

Using stochastic frontier analysis to measure the impact of weather on the efficiency of electricity distribution businesses in developing economies

Karim L. Anaya^{1,2} and Michael G. Pollitt³

Abstract

This paper analyses the influence of weather variables on the efficiency of electricity distribution companies in Argentina, Brazil, Chile and Peru, covering 82 firms which represent more than 90 per cent of the distribution market of energy delivered for the period 1998-2008. Stochastic frontier analysis (SFA) is applied using a translog input distance function. Two different approaches are evaluated: weather in the production function and weather in the inefficiency term. The efficacy of one over the other is determined using nested models. Weather data are collected from meteorological stations (429) and NASA (3,423 coordinates). A geographic information system (GIS) is used for locating the firms' service areas and their weather conditions. A combination of cost only and cost-quality models is proposed. For cost only models, the results suggest that on average there is a significant increase in measured efficiency when weather is incorporated in the production function. Under the cost-quality models, on average the effect of weather is much lower. This suggests that firms have internalised the effects of weather and have adapted their networks to the environment in which they operate. A company-level analysis indicates that across models a significant number of companies are affected by weather. Regulators are advised to make proper adjustments of efficiency scores when specific firms face important efficiency changes due to weather.

Key words: economics, technical efficiency, stochastic frontier analysis, electricity distribution markets, weather

¹ Corresponding author.

² Judge Business School, University of Cambridge, Trumpington St., CB2 1AG, Cambridge, United Kingdom. Telephone: +44(0) 1223 339700, Email: k.anaya@jbs.cam.ac.uk.

³ Judge Business School, University of Cambridge, Trumpington St., CB2 1AG, Cambridge, United Kingdom. Telephone: +44(0) 1223 339615, Email: m.pollitt@jbs.cam.ac.uk.

1. Introduction

It is important that electricity regulators are aware of the weather conditions that characterise distribution network operating service areas in order to properly evaluate the cost effect that regulated businesses could face due to unfavourable weather conditions. This is particularly true when weather conditions vary substantially within and between countries, whose electricity distribution companies might otherwise be comparable. Similarly, distribution companies need to collect and analyse weather variables and to adapt their network to the particular characteristics of the service area in order to reduce the risk of failures in their systems. The timely response to failures is a key issue for improving distribution system reliability. Thus distribution companies may want to learn appropriate lessons from other companies, in spite of actual or perceived weather differences.

The role of environmental variables or non-traditional inputs (such as weather), has been recognised in both parametric and semi-parametric stochastic frontier methods (Sun et al., 2015). In agreement with different studies, we found that environmental variables, such as those representing weather, play an important role in the efficiency of electricity distribution firms (see Table 2, Section 3). Furthermore, in the practice, weather variables have been included in the evaluation of electricity distribution networks' efficiency by energy regulators. For instance in Norway, the Norwegian Water Resources and Energy Directorate (NVE) introduced in 2007 data envelopment analysis (non parametric method) for benchmarking the distribution electricity networks. Snow and coast (which includes average wind speed it its estimation) were among the 9 output variables. In 2010 a regression stage was introduced for first time⁴. In the latest regulatory period (which started in 2016), 5 environmental factors have been included in the second stage. Some of them involve different weather variables such as snow, wind, ice and temperature (Bjørndal et al., 2016).

This paper examines how technical efficiency estimates change in the face of the inclusion or non-inclusion of weather data. A parametric stochastic frontier analysis (SFA) is proposed for evaluating the influence that weather has on firms' efficiency using a large sample of international data⁵. Technical efficiency in 82 electricity distribution firms that operate in Argentina, Brazil,

⁴ In this approach the DEA-scores (first analysis) are adjusted using a regression analysis. The new variables to be included in the second stage define the independent variables, while the DEA-scores define the dependent variables (NordREG, 2011).

⁵ We are aware that if the model proposed does not fully capture all the exogenous variables that may influence the electricity networks efficiency, then we are not looking at a perfect empirical analysis. We agree that there are other exogenous variables that may affect the electricity firms' efficiency (Migueis et al, 2012). However weather related

Chile and Peru is measured⁶. Due to the parametric nature of SFA, this method imposes restrictive assumptions related to both the production technology and the composed-error⁷. The impact of weather is evaluated across different models that have been categorised into cost only models and cost-quality models. This classification is useful in order to evaluate any kind of trade-off when quality is taken into account. We analyse the influence of weather from three perspectives: a global level, a company-level and at ownership type-level. In contrast to most other studies, our study compares the influence that weather may have on the production function and on the inefficiency term. The selection of one or other approach is analysed using nested models which help determine the approach that provides the best fit to the sample data. Firms' service areas were geo-referenced based on the country's administrative and political boundaries (e.g. department-Argentina, municipal-Brazil, commune-Chile and district-Peru). Due to the extension of the dataset (at company-level) GIS was the tool used in order to associate the weather variables with the respective companies' service area. The application of this tool can be expanded to other sectors in which geographical and environmental conditions also play an important role in efficiency analysis (i.e. health, water, transport, agriculture, etc.).

The results of this study aim to help advance understanding of the importance of including exogenous factors such as weather, due to its impact on firms' efficiency in terms of cost and quality. With the inclusion of weather in the models, this paper looks towards a more integrated approach to electricity distribution regulation when doing international comparisons. This study addresses the question within the context of stochastic frontiers, of how firms respond (in terms of efficiency change) when weather and quality issues are taken into consideration. In addition, this study identifies the countries and firms that are exposed to less favourable weather conditions and vice versa. Furthermore, this study also sheds light on the question of whether or

events are among the most important and representative exogenous variables that electric utilities need to deal with. The authors have collected and included an important set of weather variables in the analysis. Based on statistical methods for measuring their significance, a selection of the most relevant variables was made. Section 2 provides further details about the influence of weather on electricity networks.

⁶ We (and regulators) are not focussed on an absolute individual evaluation of the firms' efficiency. We are looking for a relative efficiency using SFA but taking into account specific environmental variables that firms are subject to. International benchmarking can help with this. The larger the sample, the better chance to identify best practice and improvement possibilities among firms.

⁷ Different directions of SFA have been proposed over time for relaxing some assumptions. For instance, semi-parametric and fully non-parametric SFA methods help to relaxing the distributional assumptions made on the production technology and/or composed-error (Fan et al., 1996; Kumbhakar et al., 2007; Horrace and Parmeter, 2011; Simar et al., 2016). The kernel estimator is among the most popular non-parametric regression technique used in these studies. The inclusion of environmental variables (in the production technology and/or inefficiency term) using semi-parametric techniques has also been discussed (Zhang et al., 2012; Sun and Kumbhakar, 2013; Sun et al., 2015). Other studies that make use of dynamic SFA have focused on allowing the trend in the inefficiency term to correct for past inefficiency and allowing the separation of technical efficiency from technical change (Desli et al., 2003; Tsionas, 2006; Zhang et al., 2015). Finally the integration of DEA and SFA into a unified framework of productivity analysis has also been explored (e.g. Kuosmanen and Johnson, 2010; Kuosmanen and Kortelainen, 2012).

not the firms have been able to adapt their networks to the climatic conditions that affect their respective service area. This study also innovates in the way non-environmental variables are captured (weather, service areas) being one of the first to make use of a geo-referenced system in combination with the stochastic frontier analysis. The GIS tool can be applied to different efficiency studies (with parametric and non-parametric methods).

This paper represents a development of previous studies that involve efficiency and weather. Some find that weather does influence efficiency (Jamasp *et al.*, 2012, Growitsch *et al.*, 2010, Llorca *et al.*, 2016), others find that weather does not produce any effect on efficiency (Nillesen and Pollitt, 2010; Korhonen and Syrjänen, 2003) and still others suggest that specific variables may act as proxies for the effect of weather on efficiency (Yu *et. al.*, 2009a). The large sample of international data will help us to assess whether or not these statements are true more generally or only in a sub sample (i.e. cost only models, cost-quality models, single country). We believe that this study is the first cross-country study with a focus on developing economies that evaluates the effect of weather on companies' efficiency and is a reference for future research in these economies with diverse and challenging weather conditions.

The structure of the paper is as follows. Section two discusses the influence of weather on electricity networks and the economic costs associated with the power outages. Section three summarises the results from recent studies that also evaluates the influence of weather on firms' efficiency. Section four explains the methods we use. Section five provides a description of the data collection and the selection of models. Section six discusses the results and the final section presents the conclusions.

2. The Influence of Weather on Electricity Network Performance

Electricity networks' components are susceptible to weather conditions. Power outages are the most obvious result of weather phenomena and the most costly⁸. These are the most noticeable is the case of severe weather (Keener, 1997). According to the U.S. Department of Energy (DOE, 1981), most distribution interruptions are initiated by severe weather-related interruptions in which inadequate maintenance is one of the main contributors. They find that failures in the distribution system are responsible for 80 per cent of all interruptions. Overhead lines face the strongest external factors such as weather and other exposures related to wildlife, fallen trees

⁸ Other kinds of power disturbance are related to power quality phenomena which refer to any deviations from perfect power, such as those arising from issues with voltage sags, surges, transients and harmonics (CEIDS, 2001).

etc. Lightning (flashes), high wind, extreme temperatures, snow, ice, storms and rain are among the weather variables that may affect the normal functioning of overhead lines. According to Gönen (2007), in the USA snow and ice storms are considered one of the most damaging weather events and the cause of extensive service interruptions on distribution systems, often because of tree falls on distribution circuits. Pabla (2005) states that lightning is responsible for about one-third of all faults on high voltage (HV) and distribution systems during storm days and that around 75-80 per cent of these faults are temporary. Lightning damage is one of the main concerns for many utilities because these cause the highest expense breakdown of distribution equipment. Keener (1997), states that around 90 per cent of power outages are caused by lightning during summer months, and that disruptions of electrical service throughout the year can happen due to lightning and wind damage associated with severe thunderstorm activity⁹. Short (2006) points out that fault rates (per 100 circuit mile per year) increase significantly in higher lightning areas¹⁰. In terms of the economic cost of lightning, Pinto and Almeida Pinto (2008) suggest the damage produced by lightning in Brazil is around 600m reals (c. US\$ 300m) per year, which represents 1 per cent of the electricity sector revenues, while in the USA the damage is c. US\$ 5bn per year.

Hines et al. (2008), in their study of the trend of blackouts in the USA for the period 1984-2006, find that 43.6 per cent of events were weather related¹¹. A report from Mills (2012) that uses US Department of Energy power disturbance data from the period 1992-2011, also finds a high percentage of weather related events (66 per cent by number of events, 78 per cent by number of customers affected). In addition, based on our own estimations and using the database from the U.S. Department of Energy (DOE DB), we find that the number of grid disturbances¹² from weather-related events¹³ represents around 49 per cent of the 1622 reported events, with a total of 122.7 million customers affected for the period 2006-2016¹⁴, see Figure 1.

⁹ The figures refer to the Duke's electricity distribution network from USA.

¹⁰ For instance, utilities that operate in southern US, which has a high lightning area, have a rate of 352 faults per year while those that operate in England with low lightning face a rate of 35 faults per year.

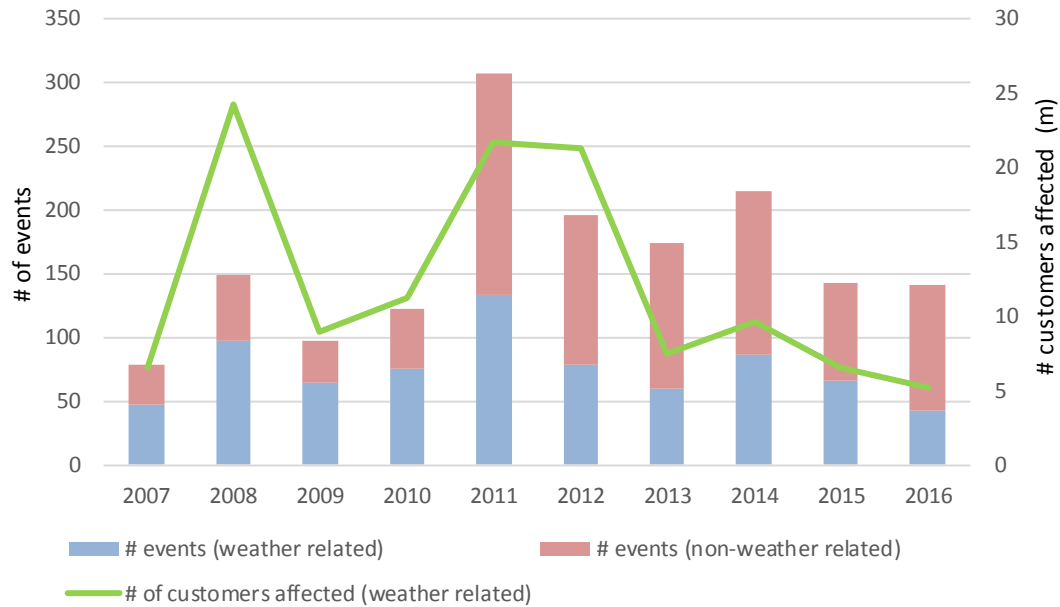
¹¹ The statistics presented in this study refer to the North American Electric Reliability Council (NERC) data from its Disturbance Analysis Working Group (DAWG). A total of 933 events that caused outages were analysed for the period 1984-2006. The events involve only those that affected more than 50,000 customers or 300 MW (demand loss).

¹² Different kinds of incidents are reported (i.e. loss of electric service, voltage reductions, islanding, load shedding, etc.), however the majority of them are associated with power outages.

¹³ Such as high winds, rain, lightning, wind storms, ice/snow storms, hail storms, thunderstorms, tornados, hurricanes.

¹⁴ Data from OE-417 Electric Emergency and Disturbance Report, Annual Summaries. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability.

Figure 1: Grid disturbance events and customers affected period 2007-2016



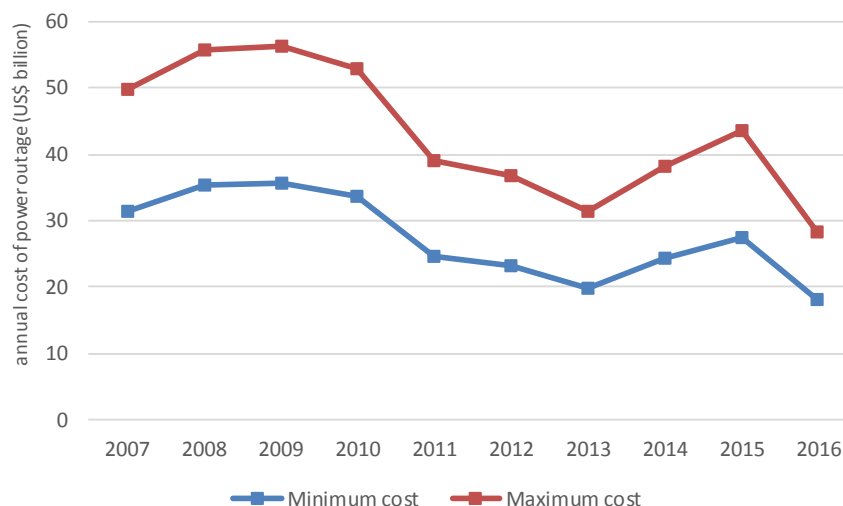
Source: U.S. Department of Energy (Forms OE-417, monthly data, annual summaries)

In general, it is observed that the number of grid disturbances from weather related events and the respective number of affected customers have decreased since 2011. However an upward trend is observed from 2007 to 2011, which is in line with Mills (2012). The economic cost of power disturbances from weather related events discussed previously is depicted in Figure 2. We have used as reference the estimations made by EPRI's Consortium for Electric Infrastructure for a Digital Society (CEIDS, 2001) regarding the economic cost of power disturbance across three U.S. economic sectors that are sensitive to power disturbances (digital economy, continuous process manufacturing and fabrication of essential services)¹⁵. These sectors account for around 2 million business establishments (17 per cent of all U.S. business establishments) representing 40 per cent of U.S. gross domestic product (GDP). CEIDS (2001) estimates that the cost of power outages across all business sectors is between US\$ 104 to US\$ 164bn a year. Based on this estimation, the assumption that 49 per cent of power outages are from weather related events (own estimation), and considering the average number of power outages with a duration over 5

¹⁵ A total of 985 business establishments were surveyed. The "direct costing" approach was used for estimating the cost of outages. This consists of asking the respondents to estimate the costs that they would incur (i.e. idle labour, materials loss, equipment damage, lost production/sales) under hypothetical outage scenarios.

minutes at 43 per cent (CEIDS, 2001, Chapter 3, p.5)¹⁶, the average annual economic cost of power outages for the period 2007-2016 would be between US\$27.4 and US\$ 43.3bn¹⁷.

Figure 2: Cost of Power Outages period 2007-2016



Source: U.S. Department of Energy (Forms OE-417), CEIDS (2001), BLS (2016)

Previous papers have also estimated the annual average cost of power outages due to weather related events. Campbell (2012) estimates these costs between US\$ 26.7 and US\$ 74.5bn. He uses two different assumptions for the percentage of power outage from weather related events, 43.6 per cent (Hines, 2008) and 78 per cent (Mills, 2012). Estimations made by the President's Council of Economic Advisers and the U.S. Department of Energy (PCEA-DOE, 2013), takes a different approach by creating a representative profile using fifteen storm profiles (occurring between 2004 and 2012) which was then applied to all the weather related power outages reported by DOE. In terms of the cost of an outage, Campbell (2012) and this paper use as reference the estimate by CEIDS (2001), while PCEA-DOE (2013) use an estimate based on Sullivan et al. (2009)¹⁸. Table 1 depicts the annual weather related outage costs related to these studies.

Table 1: Annual weather related outage costs

¹⁶ The U.S DOE defines different criteria for the registration of power disturbances events, including the size of demand loss (MW) and length of outage (with a minimum value of 15 minutes).

¹⁷ Adjusted based on CPI from the U.S. Bureau of Labor Statistics (BLS, 2016).

¹⁸ In contrast with CEIDS (2001), Sullivan et al. (2009) include in their estimation the cost of power outage not only for commercial and industrial customers but also for residential customers, which was estimated using the willingness to pay to avoid power outages. The estimates were obtained by evaluating a collection of 28 customer value of service reliability surveys performed by 10 major electric utilities from the U.S. for the period 1989-2005. Customer interruption costs were obtained by type of customer, duration of interruption, time of the day, season and day of the week.

Paper	Data (outage)	Data (cost of outage)	Period of study	Annual weather related outage costs ^{1/}	
				Min (US\$ bn)	Max (US\$ bn)
Campbell (2012)	Hines (2008), NERC DB	CEIDS (2001)	1984-2006	26.7	42.1
	Mills (2012), DOE DB	CEIDS (2001)	1992-2010	47.3	74.5
PCEA-DOE (2013)	DOE DB and DOE Emergency Situation Reports	Sullivan et al. (2009)	2003-2012	18.8	34.5
This paper (2017)	DOE DB	CEIDS (2001)	2007-2016	27.4	43.3

1/Inflation-adjusted (2016 dollars)

Based on the previous discussion we note that weather significantly and adversely influences the performance of electricity networks. Even though, this section is limited to the analysis of the U.S. power disturbance events during a specific period (due to the comprehensive database available), it helps to understand the effect that exogenous variables such as weather can produce on electricity networks. Distribution networks are quite vulnerable to bad weather conditions. Around 90 per cent of power outages occur in distribution networks. Doing the right investment helps to mitigate the impact of power outages. The upgrade of wooden poles to concrete, steel or a composite material, and the installation of associated cabling and other structural supports, can alleviate this impact (PCEA-DOE, 2013). Placing more underground lines can also help however there are many challenges. Among these are higher installation and O&M costs, higher facility replacement costs, longer repair time, more complex operational needs, among others¹⁹. Smart grid technologies (i.e. control/automatization systems, smart metering) can also enhance resiliency. For instance in the USA, the Smart Grid Investment Grant (SGIG) program represents a US\$7.9bn joint investment between the U.S. government and grant recipients²⁰. Projects fell into four categories, from which 56% of the total investments was allocated to Advanced Metering Infrastructure (AMI) and 27% to Electric Distribution System. According to the SGIG Final Report (DOE, 2016), distribution automatization and AMI have enhanced reliability with fewer and shorter outages, faster service restoration and customer services. Weather forecast improvements can also reduce the economic impact of customer outages. Economic benefits due to better weather prediction can be significant for electric utilities (Keener, 1997).

3. Previous Studies of the relationship between a Firm's Efficiency and Weather

¹⁹ New construction of underground for underground costs between US\$0.3m per mile (for rural construction) to US\$4.5m per mile (for urban construction) while converting overhead to underground distribution costs between US\$0.16m per mile (for rural construction) to US\$5m per mile (for urban construction) according to EEI (2012).

²⁰ The SGIS program has been funded by US\$3.4bn invested through the American Recovery and Reinvestment Act in order to modernise the nation's electricity system. The program involved 99 competitively selected projects, 228 utilities and organisations and was implemented between 2010 and 2015. Total SGIG investment represents 1/3 of U.S. smart grid spending (US\$24.97bn) for the period 2010-15 (DOE, 2016).

There are a number of empirical studies that try to explain the effect of environmental variables on the efficiency of electricity network companies. Studies that include in their models costs and physical, quality and environmental variables (weather and geographic conditions) are summarised in Table 2.

Table 2: Literature Review

Authors	Research	Period	Method	Environmental variables	Findings
Korhonen and Syrjänen (2003)	evaluation of cost efficiency of 102 Finish electricity distribution firms	1998	DEA. Selection of final variables based on correlations and lineal regressions	Forest cover, average snow depth	Exclusion of environmental variables due to low variation in efficiency. Few firms with high change of efficiency (10-40%). An individual analysis is recommended for some of them
Yu <i>et al.</i> (2009a)	effect of weather on the cost and quality performance of 12 UK electricity distribution utilities	1995/96 - 2002/03	Two-stage DEA and factor analysis (selection of relevant weather variables). Technical and economic efficiency	Factor 1 (min. temp, air frost, ground frost and concrete temp.), Factor 2 (total rainfall, max. temp., thunder, hail and gale)	The effect of weather on efficiency performance is small on average. Weather affects economic efficiency only under specific models, which in addition to cost and physical variables include distribution losses and customer minutes lost. Weather does affect technical efficiency if network length is dropped from the models, which suggests that network length internalises the effect of weather on efficiency scores
Growitsch <i>et al.</i> (2010)	effect of geographic and weather conditions on the cost and quality performance of 128 Norwegian electricity distribution utilities	2001-2004	Stochastic frontier analysis under the time-varying inefficiency and factor analysis (for selecting weather variables)	Around 95 variables (reduced to 7 composite factors). Include weather (precipitation, temperature, wind speeds, lightning, etc.) and geographic conditions	When comparing Battese and Coelli's 1992 and 1995 models, the study finds that the incorporation of geographic and weather variables (factors) on the inefficiency term increases the average efficiency by more than ten percentage points. When comparing Greene's 2004 and 2005 models, the study suggests that the average efficiency does not vary when geographic and weather conditions (factors) appear on the inefficiency term
Nillesen and Pollitt (2010)	the effect of error measurement and environmental factors on 109 US private electricity companies' performances	2003	DEA and Tobit regression (two-stage approach). The first layer of best-practice companies is excluded in order to correct the measurement error	Number of heating degree days (HDD) weighted by population, HDD 30 year simple state average, average 3-day maximum snowfall	Climate conditions do not explain the differences in relative efficiency. However, after doing some corrections such as comparing firms under the sample average environmental conditions, results suggest that more extreme climate factors have a negative impact on efficiency
Jamasb <i>et al.</i> (2012)	estimation of marginal cost of quality improvements in 12 UK electricity distribution utilities	1995/96 - 2002/03	Parametric approach with translog cost function	Minimum temperature (lowest monthly average), concrete temperature (number of days with temp. below 0°C), hail and thunder (number of days).	The selected weather variables are the ones that control the most of the effect of weather conditions on costs. The authors conclude that electricity distribution companies are not encouraged to improving service quality (minutes lost). Marginal costs of quality improvements exceed the incentives offered by the UK regulator
Llorca <i>et al.</i> (2014)	analyse the efficiency of 59 US electricity transmission companies taking into consideration technological differences among firms	2001-2009	Latent class model (LCM) and DEA. LMC allows to group the firms in different classes, where each class is associated to a different technology	Annual minimum temperature, average of daily precipitation and average of daily mean wind speed	LCM allocates observations into different classes better than other approaches and also predicts better the underlying efficiency of each observation. The use of clustering methods allows to capture better the underlying heterogeneity than by using non-clustering methods
Llorca <i>et al.</i> (2016)	analyse the efficiency of 59 US electricity transmission companies	2001-2009	Parametric approach with translog cost function	Annual minimum temperature, average of daily precipitation and average of daily mean wind speed	Adverse weather conditions affect negatively the transmission utilities in terms of efficiency and costs. Instead of investing in additional operating costs, investing in capital is a better strategy to handle adverse weather conditions

A number of observations on the literature can be made. First, benchmarking methodologies vary among the previous studies. Parametric and non-parametric (DEA) approaches are both popular. Second, when a number of weather variables are available, factor analysis appears to constitute a useful tool for simplifying the number of weather variables. However, it is important to take into consideration the disadvantages that this approach has. These are associated with inappropriate weights and with the difficulty of computing marginal costs for quality improvements when weather composites are included.

Benchmarking studies that include weather variables are relatively recent and limited in comparison with those where weather is omitted. There are a number of likely reasons for this. First, there has been a growing interest in climate issues and quality of supply performance in recent years. Second, the low number of studies could be associated with the difficulty of collecting weather data, especially if developing economies are the object of study. Third, all the benchmarking studies are single-country studies and are focused only on developed economies. Fourth, there is empirical evidence that weather can affect companies' efficiency. The significance of the efficiency change depends on the model specification. The inclusion or exclusion of specific variables in the model can affect the overall results. Even though some of the studies suggest that on average weather does not produce important changes in technical efficiency, individual companies can be significantly affected by weather conditions.

4. Methods

Stochastic Frontier Analysis (SFA) is our selected method for measuring technical efficiency. SFA was developed simultaneously by Meeusen and van den Broeck (1977) and Aigner *et al.* (1977). SFA allows for the incorporation of the error term that is composed of the stochastic component and the non-negative inefficiency term. SFA enables multiple inputs and outputs in the form of the distance function initially proposed by Shephard (1970). In this study, it is useful to adopt the input distance function with the translog functional form²¹. The restrictions required for the homogeneity of degree one in inputs and the symmetry assumptions for the second order coefficients are applied. Panel specification with time varying inefficiency structure is selected

²¹ The translog functional form provides a second-order differential approximation. In comparison with the linear and Cobb-Douglas functional forms which provide a first-order differential approximation, the translog functional form does not impose restrictions on the first or second derivatives itself. Its coefficients represent elasticities thus the results are interpreted quickly. The flexibility of a translog functional has a cost due to the increased number of parameters to be estimated. For further details see Christensen *et al.* (1973). Other issues can be those related to endogeneity. Kumbhakar (2013) suggests the use of instrumental variables for endogenous outputs when an input distance function is estimated.

because it is necessary to measure the trend of efficiency over time. The input distance function is normalised based on one input. The introduction of the environmental component in the production function relates to the following equation:

$$\begin{aligned}
 -\ln x_{Kit} = & \alpha_0 + \sum_{m=1}^M \gamma_m \ln y_{mit} + \frac{1}{2} \sum_{m=1}^M \sum_{n=1}^M \gamma_{mn} \ln y_{mit} \ln y_{nit} \\
 & + \sum_{k=1}^{K-1} \beta_k \ln x_{kit}^* + \frac{1}{2} \sum_{k=1}^{K-1} \sum_{l=1}^{K-1} \beta_{kl} \ln x_{kit}^* \ln x_{lit}^* \\
 & + \sum_{k=1}^{K-1} \sum_{m=1}^M \alpha_{km} \ln x_{kit}^* \ln y_{mit} + \sum_{j=1}^J \omega_j z_{jit} + \theta_t t + \frac{1}{2} \theta_{tt} t^2 \\
 & + \sum_{k=1}^{K-1} \delta_k \ln x_{kit}^* t + \sum_{m=1}^M \varphi_m \ln y_{mit} t + v_{it} - u_{it} , \\
 x^* = & \frac{x_{kit}}{x_{Kit}} \quad (1)
 \end{aligned}$$

where x_{kit} is one of the k -th input of firm i ; y_{mit} is one of the m -th output of firm i ; $\alpha, \gamma, \beta, \delta, \theta$ and φ are the parameters to be estimated; t is the time trend. z_{jit} is one of the j -th environmental variables of firm i ; v_{it} represent the stochastic component and u_{it} the non-negative inefficiency term.

Environmental variables are represented by weather variables and they have not been expressed in logs due to the existence of negatives and zero values²². Following Battese and Coelli (1992), the trend of inefficiency term over time can be represented as:

$$u_{it} = u_i \exp(-\eta(t - T)) \quad (2)$$

²² The introduction of interaction terms (weather in production) may help to give a better understanding of the relationships among the variables that have been specified in the model. However, the reason for not including interaction terms is two-fold. First, we think that this would be more appropriate in places that are subject to more extreme weather conditions that may increase the energy demand (such as in Europe and some states in the USA). In Europe (EU28, Norway, Switzerland, Iceland) demand for heating and cooling in the residential sector represented 85% of the total final energy consumption in the residential sector in 2012 and space heating alone was 78% in this sector. The residential sector alone accounted to 45% of the final energy heating and cooling consumption in 2012 (EU, 2016, p. 4, 8, 10). In New York, according to the EIA's 2009 Residential Energy Consumption Survey (EIA, 2009), heating and cooling demand represented 74% of total household site end-use energy consumption with space heating alone around 56% of this total. Countries that are part of this study, are usually not subject to extreme weather conditions. For example, in Brazil space heating is nearly non-existent (Rudnick et al., 2014) and heating and cooling (provided mainly by electric means) only represent on average 20% of overall appliances' electricity consumption (Tubelo et al., 2014). Second, weather per se is a complex variable, thus we find it convenient to keep this simple and to avoid the use of interaction terms that involve weather. This is in line with other studies that have evaluated the influence of weather on electricity firms (Growitsch et al., 2010; Nillesen and Pollitt, 2010; Jamasb et al., 2012; Llorca et al., 2014, 2016).

where η (eta) is the unknown parameter to be estimated, T is the last time period of i -th firm, u_i is associated with technical inefficiency, is independent and identically distributed and has a truncated normal distribution, $N^+(u, \sigma_u^2)$.

The η parameter allocates a common technical efficiency trend among producers. This is one of the main disadvantages of this approach²³. When η is higher than 0, that means $\eta(T - t) > 0$ technical efficiency improves over time ($u_{it} > u_i$). Similarly, when η is lower than 0 technical efficiency decreases over time and when η is equal to 0, technical efficiency does not vary over time. It is assumed that environmental variables influence the shape of the input distance function directly. A different approach is that in which environmental factors are included in the inefficiency term (Kumbhakar *et al.*, 1991). In this case, the maximum likelihood estimates would be computed under the assumption that inefficiency has a distribution that varies with Z and that is no longer identically distributed. Thus, u_i would be defined as follows:

$$u_i \sim N^+(s_i, \sigma_u^2) \quad (3)$$

where, $s_i = \omega_o + \sum_{j=1}^J \omega_j Z_{ij}$ and ω_o, ω_j are the parameters to be estimated.

Battese and Coelli (1995) propose a similar model but applied to a panel data context. For a better discussion of the results, this last approach has been also applied in order to compare both methods (environment in production function versus environment in inefficiency term). The use of specific tests (such as log likelihood ratio test) is appropriate for determining the approach that provides the best fit. The treatment of environmental variables has been discussed in different studies. In summary, it is clear that some of them assume that the environment can affect the shape of the production function (in this case input distance function). Other studies support

²³ Cuesta (2000) proposes a modified version of this approach in which a common pattern of inefficiency change is not imposed and the pattern can be tested against firm-specific patterns. The equation proposed is as follows: $u_{it} = u_i \exp(-\varepsilon_i(t - T))$, where ε_i are firm-specific parameters that respond to different patterns of temporal variation among firms. Gauss was the tool selected for estimations. He applies this method to Spanish dairy firms. Even though Cuesta (2000) represents a more flexible approach by not imposing a specific temporal pattern on efficiency, companies that operate in countries from the same region and are in the same industry may be subject to a common temporal pattern. Our sample consists of companies that operate in developing economies in the South American region. Our approach in line with Lee (2010) who uses different methods (including the one from Cuesta (2000)) for evaluating the growth and efficiency of 49 countries over the period 1965-90 based on the Penn World data. The countries were divided into four groups, including one group composed of countries from South America, Middle East and Africa. He shows that the four groups have different temporal patterns (five hypotheses were tested for this purpose, all of them rejected at the 1 per cent level). Lee (2010) also suggests that the Cuesta (2000) approach may conduce to imprecise estimates of the production function when the panel data sample has a large number of firms and the time series observation per firm is small (as our sample: 82 firms over 11 years).

the idea that environmental variables act as explanatory variables of inefficiency only²⁴ (i.e. in the second stage of a two-stage approach). The first step of the two-stage approach consists of estimating the conventional frontier model but omits the environmental variables. In the second step, the predicted technical efficiencies are regressed on a set of explanatory variables (environmental variables). Some authors find important econometric problems when applying the second stage approach (e.g. Simar and Wilson, 2007). First, it is assumed that the exogenous variables are not correlated with the remaining independent variables. If this is the case, the estimates are biased because of the omission of the exogenous variables in the first stage stochastic model. Thus, the predicted efficiencies in the second stage are biased as well. Second, inefficiencies are identically distributed in the first stage, however, in the second stage a functional relationship is assumed between predicted efficiencies and environmental variables, see Kumbhakar and Lovell (2000) and Coelli *et al.* (2005). Simar and Wilson (2007) provide a detailed explanation of the second stage approach and the ways to improve their estimates using single and double bootstraps. Additional approaches to environmental variables are also discussed in some frontier efficiency studies, involving third and fourth stages. See Yang and Pollitt (2009) for a complete description of the different stages and their application in the context of measuring efficiency in a sample of Chinese coal-fired power plants.

5. Data Collection and Models

Our data consists of an unbalanced panel for 82 electricity distribution companies for the period 1998-2008. The companies operate in Argentina (18), Brazil (39), Chile (11) and Peru (14) and account for more than 90 per cent of the total distribution market in those four countries in terms of energy delivered. Among the data that was collected are (1) cost data, (2) physical data, (3) quality variables and (4) weather variables. The main sources of information are companies' annual reports and many of them were collected during in-country fieldwork from energy regulators, associations of electricity distribution companies, the companies themselves and from the national securities market commissions and insurance agencies. Information obtained from system operators and from the World Bank supplemented the data on network length (World Bank, 2008). Weather data was collected from meteorological offices and NASA. Table 3 shows the 2008 descriptive statistics for the 82 companies across the countries.

²⁴ See Mota (2004). Coelli *et al.* (1999) and Growitsch *et al.* (2010) compare their results under the following approaches: environment in production and environment in inefficiency term.

Table 3: 2008 Descriptive Statistics – Distribution Electricity Companies

Variables (2008)	Units	Argentina			Brazil			Chile			Peru			Total		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Cost variables																
OPEX	US\$ (million)	15.3	375.8	115.7	14.9	1010.0	191.0	12.86	269.3	73.8	3.1	127.7	34.4	3.1	1010.0	132.31
CAPEX	US\$ (million)	9.45	228.34	55.68	4.7	583.8	168.8	3.32	175.5	41.1	2.5	122.1	35.4	2.5	583.8	106.40
Physical variables																
Customers (residential)	million	0.09	2.24	0.49	0.08	5.48	1.40				0.02	0.97	0.28	0.02	5.48	0.942
Customers (total)	million	0.10	2.53	0.66	0.10	6.69	1.64	0.06	1.53	0.44	0.03	1.03	0.33	0.03	6.69	1.1
Energy delivered (residential)	GWh	252	7738	1353	160	14427	2468				23	2066	445	23	14427	1777
Energy delivered (free market) ^{1/}	GWh	40	3700	691	0	19459	2185	12	2687	536	0	929.05	118	0	19459	1178
Energy delivered (cooperatives)	GWh	0	1728	236										0	1728	236
Energy delivered (total)	GWh	672	18616	4891	492	41898	9126	218	12535	2545	82	5334	1159	82	41898	5982
Length of network	Km (000')	6.0	51.8	17.7	2.9	453.4	75.3	1.0	26.9	11.2	0.9	21.5	10.6	0.9	453.4	43.8
N° of workers	number	85	3130	891	153	8031	1798	57	790	322	26	666	270	26	8031	1132
Service Area	Km ² (000')	3.3	203.0	86.9	1.8	1570.8	248.0	2.1	87.4	44.7	2.4	140.9	45.0	1.8	1570.8	72.7
Quality variables																
Total losses	GWh	57	2247	718	44	6686	1437	20	786	229	9	510	124	9	6686	897
Total losses	percentage	7.8%	23.8%	12.3%	4.5%	37.2%	15.6%	5.9%	14.7%	9.6%	8.1%	14.1%	11.1%	4.5%	37.2%	13.3%
Technical losses	percentage				3.6%	23.7%	8.7%							3.6%	23.7%	8.7%
Non-technical losses	percentage				0.2%	25.7%	5.4%							0.2%	25.7%	5.4%
Average Frequency of interruptions/customer	times/year	5.5	12.5	8.8	5.2	51.6	14.7				5.8	36.5	18.4	5.2	51.6	14.8
Average Duration of interruptions /customer	hours/year	8.7	21.5	14.9	5.9	77.2	19.9				13.4	148.0	42.2	5.9	148.0	24.3
Other Variables																
Private companies	number		Total			Total			Total			Total			Total	
			16			29			11			4			60	
Public companies	number		2			10			0			10			22	
Total companies	number		18			39			11			14			82	

^{1/} In the case of Chile it refers to 2007 figures.

Sources: Companies' annual reports, National and Regional Energy Regulators, Energy Ministries, Associations of Electricity Distribution Utilities

5.1 Cost, Physical and Quality Data

Operating costs (opex) are composed of (1) distribution cost, (2) retail cost and (3) administrative and general expenses. The way of presenting financial figures was not homogenous across companies and consequently national and regulatory accounting was analysed for grouping cost figures based on three sub categories: labour cost, materials and third party services (ANEEL, 2007; SEC, 2006; MINEM, 1994). Costs associated with the transmission and generation business were excluded. Capital costs (capex) are represented by total asset additions, including work in progress. All figures were adjusted to 2008 prices using the consumer price index (CPI) and the purchasing power parity (PPP) exchange rate as appropriate.

In terms of physical data, the number of customers is composed of residential, industrial and rural businesses as well as government customers. Energy delivered refers to the total sale of energy in the regulated and the free market (i.e. industrial and larger commercial customers). In the case of Argentina, it also includes energy delivered to cooperatives. Length of network refers to the distribution business; however this concept varies across countries. Based on an individual analysis among countries, this study concludes that distribution networks in general are those with voltage levels up to 34.5 kV, that are associated with low and medium voltages. In most of the cases, higher voltages are part of the transmission business and have been excluded.²⁵ Service area represents the area in which the companies operate. Some specific reports and databases were used in order to locate the service area geographically for each firm using a geographic information system (GIS). Further details are given in section 4.2.

Power losses are composed of technical and non-technical losses²⁶. Interruptions involve those that are equal to three minutes or longer, both planned and unplanned, internal and external but exclude major interruption events. In the case of Argentina and Peru, quality variables regarding the duration and frequency of interruptions were provided directly from regulators in the format

²⁵ For instance, transmission lines represent around 4.5 per cent and 2.5 per cent of the total length regarding electricity distribution firms from Argentina and Chile respectively. 2008 Figures. It is also noted that in Brazil electricity distribution firms may operate in three different categories: 220/127 V, 380/220 V, 3.8 KV, 13.8 KV, 15 KV, 20 KV, 23 KV and 34.5 KV for distribution; 69 KV and 88 KV for sub transmission and 138 KV and 230 KV for transmission. Bearing in mind that the majority of distribution is concentrated in the first category, and in order to be in line with the criterion applied for the rest of companies, it is prudent to associate length of network with the first category only (up to 34.5 kV).

²⁶ It is noted that a better approach would be to include only technical losses, however it was not possible to split the total losses into these two components (excluding Brazil). Thus, total losses have been adopted just as a proxy for representing the effect of including quality variables, influencing the behaviour of each company and then the efficiency of distribution firms.

required. In the case of Peru, it was not possible to exclude major interruptions and the interruptions refer to those produced in urban areas. However around 80 per cent of the total number of customers are concentrated in these areas (OSINERGMIN, 2003).

5.2 Weather Data

Weather data were collected from national meteorological offices²⁷ and from the National Aeronautics and Space Administration of the USA (NASA). Table 4 summarises the weather data per country and type of variable. The meteorological offices provided information regarding weather data that was recorded in 458 stations, of which 429 are placed inside the service area of the companies that are part of this study. Meteorological offices provided coordinates (latitude, longitude) for each met station. ArcGIS is the application that was used to match this data to individual firm service territories.

Maximum absolute temperature; minimum absolute temperature; total rainfall; number of days in a year with: gales, storm, hails and frost days; and humidity are among the data collected (monthly data). NASA provided lightning data (number of flashes/km²/year). The data set used was that from LIS (Lightning Imaging Sensor) HRFC (High Resolution Full Climatology), with tropical coverage for the period 1998-2008 with a resolution of 0.5 degrees. Similar to the procedure followed previously, a geographic information system was used for plotting the flash rate coordinates. Around 3,423 coordinates (grid data) with information about flash rates were identified inside the service area of the whole sample of companies²⁸. The location of the companies' service area in a geo-referenced system was required for matching the meteorological stations and flash rate coordinates for each firm. The first step was to get the digital maps for the four countries. Depending on the country's administrative and political boundaries, maps could be obtained at the level of the department (Argentina), municipal (Brazil), district (Peru) and commune (Chile). Usually these boundaries are related to the service or concession area that is allocated to a specific utility.

²⁷ National Meteorological Service (SMN) of Argentina, National Institute of Meteorology (INMET) of Brazil, Meteorological Direction (DMC) of Chile and the National Service of Meteorology and Hydrology (SENAMHI) of Peru.

²⁸ Based on the coverage (~35°N/S) the total number of coordinates provided by NASA is 100,800 (720*360) per year. Gauss was the software used for arranging the data in the format required for ArcGIS.

Table 4: 2008 Descriptive Statistics – Weather Variables

Variable (2008) ^{1/}			Units			Argentina			Brazil			Chile			Peru					
						Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean			
Temperature ^{2/}																				
Max. Absolute	tmax	degrees Celsius	37.0	42.8	40.2	34.0	42.1	38.5	27.5	37.3	33.5	22.0	39.2	33.6						
Min. Absolute	tmin	degrees Celsius	-18.8	-1.8	-6.2	-5.4	21.2	7.2	-6.6	9.9	-0.7	-13.5	16.8	2.3						
Rain ^{3/}																				
Total rain	rain	mm	208.5	1060.6	657.4	822.1	2782.8	1560.0	0.0	1538.2	486.2	5.6	2139.4	542.1						
Wind																				
Gales ^{4/}	gal	Nº days/year	0.0	36.0	14.5							0.0	1.5	0.2						
Max. Speed	ms	Km/h year				5.6	17.0	9.9												
Humidity (relative)																				
Average	hum	percentage				61.63	86.8	74.7	54.8	80.4	69.5	59.5	83.7	71.6						
Maximum	humax	percentage				96.9	100.0	99.1	99.4	99.4	86.7									
Flashes ^{5/}																				
Nº flashes per km	fr	flashes/km ² /year	2.0	7.0	4.2	0.6	8.2	4.2	0.0	0.7	0.2	0.0	4.8	1.4						
Others																				
Storms	st	Nº days/year	18.0	81.0	38.0															
Hails	hail	Nº days/year	0.0	2.5	1.0															
Frost days ^{6/}	fd	Nº days/year	4.0	64.3	21.0							0.0	64.0	17.2	0.0	145.0	20.1			
Met Stations / NASA																				
Coordinates						Total			Total			Total			Total					
Total Met Stations		number	46			293			16			74								
Total NASA Coordinates		number	408			2704			108			203								

^{1/} All figures refer to the minimum, maximum of mean value inside a specific company' service area. For instance, in the case of Argentina, the minimum value of rain refers to ESJ and the maximum value refers to EDESA .

^{2/} Based on average monthly data.

^{3/} Rain is defined as total annual value.

^{4/} Gales are defined as those winds with speed equal to 63km/h or higher. In the case of Peru, figures refer to 2007 period. Data were provided only for the period 1998-2007.

^{5/} Flashes refer to the number of lightnings per square km.

^{6/} Frost days are those days in which the minimum air temperature falls below 0 C°. This variable is also known as "helada".

Source: SMN, INMET, DMC, SENAMHI, NASA

The digital maps were obtained from the National Agricultural Research Centre (INTA) from Argentina, the geo-referenced information system of the electric sector (SIGEL) from Brazil, the National System of Coordination of Territorial Information (SNIT) from the Ministry of National Property of Chile, and from the National Geographic Institute (IGN) from Peru²⁹. The second step was to get the companies' service area data set. The information at the departmental level was found in the annual reports from the Secretary of Energy from Argentina (Secretaria de Energia, 2008). In the case of Brazil, the digital map includes this dataset. The Chilean data set was provided by the Superintendence of Electricity and Fuel (SEC). The Peruvian data was provided by OSINERGMIN and the National Institute of Statistics and Information (INEI). With the digital maps and firms' service area data set, it was possible to geo-reference the firms' service areas and allocate the meteorological stations and flash rate coordinates for each one. The number of meteorological stations and flash rate coordinates associated with a firm's service area was higher than 1. Thus, averages were taken. On average we have the following ratios: 3.3 for Argentina, 7.5 Brazil, 1.5 Chile and 5.3 Peru respectively. The maximum ratio of stations per firm's service area is 51 (CEMIG, a Brazilian company)³⁰.

5.3 Selection of Variables and Model Specifications

In this section the selection of the preferred variables and models (cost, physical, quality and weather variables) is discussed. Six models have been selected, see Table 5. These variables have been selected based on previous studies³¹. This study starts by assuming that weather has a direct influence on the production function and that each utility faces a different production frontier. The other option is to consider that weather variables influence the inefficiency term directly, which means that weather would impact only on the difference given by the deviations from the frontier. We compare the two approaches in section 5.1. Among the studies that add environmental variables to the production function are Pollitt (1995), Estache *et al.* (2004), Rossi (2007) and Jamasb *et al.* (2012). By contrast, Mota (2004), Nillesen and Pollitt (2010) and Growitsch *et al.* (2010) assume that the environment influences directly on

²⁹ 535 departments (Argentina), 5,562 municipalities, (Brazil), 342 communes (Chile), 1,833 districts (Peru).

³⁰ We are aware of the aggregation issues that arise when only the average value of weather variables is taken to represent the effect of weather in a specific electricity firm (i.e. depending on the size of the firm, its service area may involve different regions such as coast and mountains, then it is exposed to different weather conditions). A better approach would have been to geo-reference each firm's distribution network and to associate the respective weather variable, however information of the network deployment was not possible to obtain (due to confidentiality issues).

³¹ Number of customers and energy delivered are the most common output variables. Regarding length of network, the consensus on its inclusion as output or input is much lower. Some studies use it as an output (proxying coverage) - see Mota (2004), Giannakis *et al.* (2005), Yu *et al.* (2009a,b), Jamasb *et al.* (2012) and Llorca *et al.* (2014); while others use it as an input -see Estache *et al.* (2004), Rossi (2007) and Tovar *et al.* (2011). In this study the first one is used. Total losses and customer hours lost are variables that firms attempt to reduce and that are usually seen as inputs, see Giannakis *et al.* (2005), Yu *et al.* (2009a,b) and Jamasb *et al.* (2012).

efficiency. Coelli *et al.* (1999) compare and discuss both approaches but in relation to the airline market.

Table 5: Models

Variable	Type of variable	Cost models		Cost-quality models			
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
OPEX (x1)	Monetary	I	I	I	I	I	I
CAPEX (x2)			I			I	I
CUST (y1)	Physical	O	O	O	O	O	O
ENG (y2)		O	O	O	O	O	O
LEN (y3)		O	O	O	O	O	O
LOSS (x3)	Quality			I	I	I	I
CHL (x4)					I		I
W1 (rain)	Weather	E	E	E	E	E	E
W2 (tmax)		E	E	E	E	E	E
W3 (tmin)		E	E	E	E	E	E

I: input, O: output, E: environment, OPEX: operating costs, CAPEX: capital costs, CUST: Number of customers, ENG: energy delivered, LEN: length of network, LOSS: power losses, CHL: customer hours lost, W1: total rainfall, W2: maximum absolute temperature, W3: minimum absolute temperature

5.3.1 Weather Variables

A total of five weather variables have been analysed: total rainfall, maximum absolute temperature, minimum absolute temperature, humidity and flash rate (lightning)³². However, the last two are not included in the reported analysis. This is because the introduction of humidity in the models does not produce any effect on the production function; its value is very weak and it is not statistically significant. We did however proceed with the analysis of the impact of flash rate, due to its apparently stronger relationship with rainfall (which appeared to indicate that either flash rate or rainfall could be used to model a distinct weather impact). Due to this fact, three scenarios per model were analysed, see Table 6.

Table 6: Weather scenarios

Scenarios	total rain (rain)	max. absolute temp. (tmax)	min. absolute temp. (tmin)	flash rate (fr)
Scenario 1	x	x	x	x
Scenario 2	x	x	x	
Scenario 3		x	x	x

³² This is in agreement with previous studies (Yu *et al.*, 2009a; Growitsch *et al.*, 2010; Jamasb *et al.*, 2012; Llorca *et al.*, 2014, 2016) and with the availability of weather data across countries.

Appendix 1 presents the parameters of weather variables when these are included in the production function and provides further details of their significance on the production function. We conclude that maximum and minimum absolute temperatures are the weather variables that most influence the shape of the technology under cost-quality and cost only models respectively and flash rate is basically not significant across the three scenarios³³. Thus, Scenario 2 was selected as the preferred model. This combines total rainfall, and maximum and minimum absolute temperature. This is the model that captures most of the effect of weather variables on the production function. This is in line with Jamasb *et al.* (2012), who suggest that given the complexity of weather variables it is better to focus on the overall effect rather than the individual effect of a specific weather variable, due to the possible correlations that can exist between weather variables.

6. Results

6.1 Maximum Likelihood estimation

The results are presented based on three cases: translog without weather (Case A)³⁴, weather in the distance function (Case B) and weather in inefficiency (Case C). Case A refers to those models in which weather has not been included, i.e. Eq.1 excluding z_{jit} and equation (2). Case B comprises the models where weather variables are included in the production function, based on equation (1) and (2). Case C is composed of models in which it is assumed that weather influences the inefficiency term, based on equation (1) excluding z_{jit} and equation (3).

This section discusses the maximum likelihood estimates for each case and supports the selection of Case B models, which is weather in the production function (vs weather in the inefficiency term). Table 7 shows the maximum likelihood estimates for all the models³⁵. STATA was used for computing Case A and Case B models and Frontier 4.1 for estimating Case C models. The variables are normalised by dividing them by their respective sample geometric means, thus first order coefficients represent elasticities at the sample mean. The time trend was also adjusted to the mean, where first order coefficients refer to the technical change at the sample mean. It is evident that the input distance function is well specified and most parameters are statistically significant. In general, first order output estimates have the correct sign, which means that the coefficients on number of customers, energy delivered and length

³³ It is important to note that at country-level, an in special in the case of Brazil lightning can be an important weather variable that may affect the operation of the distribution network.

³⁴ In this scenario, weather variables have been removed across the six models.

³⁵ All the regression coefficients have been rounded to 3 decimal places. In terms of weather variables, coefficients z_1 (rainfall) are the lowest ones but different from zero.

of network are negative. Some exceptions are observed but these coefficients are not statistically significant.

Output elasticities for Case A, Case B and Case C models sum on average to -0.9648, -0.9593 and -0.9590 respectively. Elasticities from cost only models are the lowest. This suggests that there are very slight increasing returns to scale at the sample mean across the three cases. The introduction of weather does not produce an important increase in economies of scale (estimated as the inverse of elasticity).

In terms of technical change in Case A (without weather variables) and Case B models, all the time coefficients are positive and statistically significant, which indicates a mean technical progress of 1.33 per cent and 1.48 per cent per year respectively. This means that the introduction of weather produces a minor upward increase of 0.15 percentage points of technical progress. The cost only models are those that contribute the most. Regarding Case C, technical change is not statistically significant except for M4 Case C. In terms of the non-neutral technical change, which is denoted by the time interacted with each output (in logs), those that correspond to customers and length of network in general have a positive impact on opex reduction, while that corresponding to energy delivered has the inverse effect. However, the coefficients of cost quality models are in general those that are statistically significant.

Regarding models that include weather variables (Case B and Case C), the coefficients have the correct sign. As previously mentioned, in Case B models, maximum and minimum temperatures are the variables that influence the production function the most. The influence of total rainfall is weak and is only statistically significant in M3 Case B. For Case C models we observe that the level of significance increases in comparison with Case B models and that the impact of total rainfall still remains weak. Furthermore, the influence of maximum and minimum temperature on inefficiency is higher in cost-quality models than in cost only models. The positive sign of these estimates indicates that inefficiency increases when these values also rise.

Table 7: Input distance function maximum likelihood estimates

Variables	Model 1						Model 2						Model 3					
	M1 Case A		M1 Case B		M1 Case C		M2 Case A		M2 Case B		M2 Case C		M3 Case A		M3 Case B		M3 Case C	
	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat	Coef.	t Stat
α_0	0.600 (7.85)		0.541 (10.29)		0.407 (9.61)		0.594 (6.45)		0.550 (8.97)		0.439 (4.49)		0.490 (10.36)		0.477 (13.56)		0.333 (15.81)	
$\ln(y_1)$	-0.525 (-7.29)		-0.538 (-7.27)		-0.721 (-17.52)		-0.506 (-7.23)		-0.518 (-7.20)		-0.671 (-15.39)		-0.596 (-10.37)		-0.604 (-10.81)		-0.629 (-20.95)	
$\ln(y_2)$	-0.318 (-6.97)		-0.294 (-6.07)		-0.217 (-8.32)		-0.329 (-7.95)		-0.307 (-6.85)		-0.238 (-9.14)		-0.313 (-9.18)		-0.287 (-8.13)		-0.319 (-15.63)	
$\ln(y_3)$	-0.098 (-2.49)		-0.099 (-2.73)		0.014 (0.64)		-0.126 (-3.13)		-0.128 (-3.50)		-0.036 (-1.71)		-0.081 (-2.79)		-0.089 (-2.96)		-0.022 (-1.36)	
$0.5*\ln(y_1)^2$	0.005 (0.02)		0.032 (0.14)		-0.232 (-1.45)		0.226 (1.00)		0.248 (1.13)		-0.415 (-3.01)		-0.475 (-2.55)		-0.380 (-2.01)		-0.357 (-2.65)	
$0.5*\ln(y_2)^2$	-0.127 (-2.81)		-0.129 (-2.88)		0.015 (0.26)		-0.136 (-3.18)		-0.138 (-3.25)		-0.003 (-0.06)		-0.110 (-3.28)		-0.119 (-3.50)		-0.069 (-1.45)	
$0.5*\ln(y_3)^2$	0.105 (1.15)		0.089 (1.04)		-0.372 (-6.77)		0.098 (1.16)		0.085 (1.05)		-0.383 (-7.69)		0.012 (0.18)		-0.008 (-0.13)		-0.153 (-3.09)	
$\ln(y_1)*\ln(y_2)$	0.076 (0.83)		0.054 (0.60)		0.031 (0.37)		0.011 (0.12)		-0.010 (-0.12)		0.097 (1.35)		0.179 (2.56)		0.119 (1.65)		0.148 (2.09)	
$\ln(y_1)*\ln(y_3)$	-0.027 (-0.20)		-0.019 (-0.14)		0.327 (3.77)		-0.137 (-1.08)		-0.121 (-0.97)		0.405 (5.16)		0.172 (1.62)		0.140 (1.33)		0.224 (2.98)	
$\ln(y_2)*\ln(y_3)$	-0.066 (-0.90)		-0.052 (-0.74)		-0.066 (-1.47)		0.015 (0.21)		0.020 (0.30)		-0.102 (-2.49)		-0.111 (-2.03)		-0.050 (-0.85)		-0.091 (-2.45)	
$\ln(x_2/x_1)$							0.139 (10.64)		0.132 (10.12)		0.218 (15.05)							
$\ln(x_3/x_1)$													0.494 (24.02)		0.487 (23.87)		0.323 (19.57)	
$\ln(x_4/x_1)$																		
$0.5*\ln(x_2/x_1)^2$							0.056 (3.16)		0.056 (3.23)		0.014 (0.64)							
$0.5*\ln(x_3/x_1)^2$													-0.231 (-5.15)		-0.218 (-4.95)		-0.344 (-6.90)	
$0.5*\ln(x_4/x_1)^2$																		
$\ln(x_2/x_1)*\ln(x_3/x_1)$																		
$\ln(x_2/x_1)*\ln(x_4/x_1)$																		
$\ln(x_3/x_1)*\ln(x_4/x_1)$																		
$\ln(y_1)*\ln(x_2/x_1)$							0.099 (2.66)		0.121 (3.32)		0.011 (0.23)							
$\ln(y_1)*\ln(x_3/x_1)$													-0.058 (-1.00)		-0.025 (-0.42)		-0.165 (-3.13)	
$\ln(y_1)*\ln(x_4/x_1)$																		
$\ln(y_2)*\ln(x_2/x_1)$							0.035 (1.75)		0.020 (1.02)		0.124 (4.88)							
$\ln(y_2)*\ln(x_3/x_1)$													0.219 (5.93)		0.192 (5.29)		0.215 (7.01)	
$\ln(y_2)*\ln(x_4/x_1)$																		
$\ln(y_3)*\ln(x_2/x_1)$							-0.099 (-4.45)		-0.102 (-4.70)		-0.076 (-2.74)							
$\ln(y_3)*\ln(x_3/x_1)$													-0.073 (-2.15)		-0.074 (-2.22)		0.006 (0.20)	
$\ln(y_3)*\ln(x_4/x_1)$																		
t	0.023 (4.85)		0.023 (5.46)		0.003 (1.03)		0.017 (3.31)		0.018 (3.77)		-0.001 (-0.34)		0.009 (2.94)		0.011 (3.62)		0.002 (0.79)	
$0.5*t^2$	-0.008 (-5.69)		-0.008 (-5.74)		-0.008 (-3.70)		-0.013 (-8.88)		-0.012 (-8.71)		-0.017 (-8.51)		-0.007 (-7.00)		-0.007 (-7.24)		-0.005 (-3.24)	
$t*\ln(y_1)$	0.000 (0.00)		-0.006 (-0.85)		0.000 (-0.02)		-0.005 (-0.74)		-0.012 (-1.73)		0.001 (0.06)		0.020 (4.01)		0.017 (3.21)		0.013 (1.64)	
$t*\ln(y_2)$	-0.002 (-0.45)		0.003 (0.68)		-0.016 (-2.46)		0.003 (0.56)		0.008 (1.76)		-0.009 (-1.50)		-0.026 (-7.39)		-0.023 (-5.98)		-0.028 (-5.46)	
$t*\ln(y_3)$	0.009 (2.22)		0.008 (2.19)		0.025 (4.47)		0.003 (0.90)		0.004 (1.01)		0.010 (1.90)		0.006 (2.14)		0.006 (2.06)		0.017 (4.15)	
$t*\ln(x_2/x_1)$							0 (-0.08)		-0.002 (-0.57)		0.004 (1.00)							
$t*\ln(x_3/x_1)$													0.024 (5.91)		0.024 (6.03)		0.020 (3.58)	
$t*\ln(x_4/x_1)$																		
z1 (rain)			0.000 (-1.09)						0.000 (-0.54)						0.000 (-1.71)			
z2 (tmax)			-0.004 (-0.87)						-0.005 (-1.07)						-0.009 (-2.62)			
z3 (tmin)			-0.007 (-2.96)						-0.005 (-2.31)						-0.005 (-2.16)			
δ_0					0.135 (1.13)						0.248 (1.61)						-0.242 (-1.33)	
w1 (rain)					0.000 (2.82)						0.000 (1.38)						0.000 (3.44)	
w2 (tmax)					0.019 (2.24)						0.008 (1.34)						0.066 (3.53)	
w3 (tmin)					0.002 (0.66)						0.007 (3.41)						0.015 (3.68)	
γ	0.819 (19.52)		0.844 (21.16)		0.831 (16.58)		0.784 (17.59)		0.804 (18.61)		0.742 (7.24)		0.959 (44.73)		0.950 (42.45)		0.911 (38.37)	
LLF	239.6		257.0703		-23.27		297.3		313.16		85.34		479.5		482.8		128.20	
N ^o of observations	809		790		790		797		788		788		807		788		788	

Case A: without weather, Case B: weather in production function and Case C: weather in inefficiency

Table 7: Input distance function maximum likelihood estimates (*continued*)

Variables	Model 4			Model 5			Model 6		
	M4 Case A Coef. t Stat	M4 Case B Coef. t Stat	M4 Case C Coef. t Stat	M5 Case A Coef. t Stat	M5 Case B Coef. t Stat	M5 Case C Coef. t Stat	M6 Case A Coef. t Stat	M6 Case B Coef. t Stat	M6 Case C Coef. t Stat
α_0	0.449 (13.21)	0.500 (12.26)	0.364 (11.02)	0.455 (11.33)	0.466 (13.27)	0.287 (12.95)	0.466 (13.21)	0.515 (12.75)	0.308 (5.17)
$\ln(y_1)$	-0.651 (-10.12)	-0.654 (-10.03)	-0.774 (-19.26)	-0.573 (-11.16)	-0.585 (-10.98)	-0.608 (-22.00)	-0.631 (-9.45)	-0.640 (-9.47)	-0.747 (-18.67)
$\ln(y_2)$	-0.314 (-7.78)	-0.280 (-6.46)	-0.196 (-6.23)	-0.325 (-10.13)	-0.295 (-8.39)	-0.325 (-17.32)	-0.309 (-7.52)	-0.270 (-6.18)	-0.235 (-7.82)
$\ln(y_3)$	0.009 (0.26)	-0.023 (-0.67)	0.013 (0.48)	-0.089 (-3.40)	-0.103 (-3.80)	-0.055 (-3.55)	-0.014 (-0.40)	-0.042 (-1.23)	0.012 (0.44)
$0.5*\ln(y_1)^2$	0.548 (1.86)	0.488 (1.65)	0.067 (0.31)	-0.339 (-1.74)	-0.272 (-1.48)	-0.453 (-4.03)	0.712 (2.35)	0.700 (2.29)	-0.137 (-0.51)
$0.5*\ln(y_2)^2$	0.245 (2.29)	0.181 (1.70)	0.204 (1.91)	-0.119 (-3.70)	-0.128 (-3.90)	-0.084 (-2.29)	0.232 (2.10)	0.185 (1.72)	0.047 (0.41)
$0.5*\ln(y_3)^2$	-0.144 (-1.67)	-0.184 (-2.23)	-0.448 (-6.40)	-0.045 (-0.72)	-0.060 (-0.98)	-0.168 (-3.57)	-0.126 (-1.52)	-0.160 (-1.99)	-0.402 (-4.31)
$\ln(y_1)*\ln(y_2)$	-0.457 (-2.92)	-0.414 (-2.59)	-0.312 (-2.12)	0.138 (2.08)	0.083 (1.21)	0.187 (3.36)	-0.484 (-2.93)	-0.474 (-2.84)	-0.116 (-0.72)
$\ln(y_1)*\ln(y_3)$	-0.073 (-0.53)	-0.059 (-0.44)	0.290 (3.28)	0.142 (1.38)	0.119 (1.20)	0.264 (3.87)	-0.129 (-0.96)	-0.130 (-0.98)	0.295 (2.53)
$\ln(y_2)*\ln(y_3)$	0.181 (2.81)	0.205 (3.11)	0.114 (1.91)	-0.062 (-1.21)	-0.008 (-0.15)	-0.099 (-2.95)	0.185 (2.83)	0.221 (3.26)	0.074 (1.21)
$\ln(x_2/x_1)$				0.084 (8.01)	0.078 (7.57)	0.178 (14.03)	0.057 (4.77)	0.055 (4.68)	0.122 (4.98)
$\ln(x_3/x_1)$	0.466 (16.73)	0.449 (15.60)	0.300 (10.33)	0.458 (22.25)	0.454 (22.55)	0.283 (16.00)	0.429 (15.19)	0.411 (14.07)	0.293 (10.76)
$\ln(x_4/x_1)$	0.092 (5.22)	0.101 (5.63)	0.088 (4.07)				0.093 (5.31)	0.102 (5.71)	0.053 (2.47)
$0.5*\ln(x_2/x_1)^2$				0.048 (3.69)	0.049 (3.83)	-0.010 (-0.56)	0.031 (1.81)	0.037 (2.19)	-0.025 (-0.82)
$0.5*\ln(x_3/x_1)^2$	-0.158 (-2.55)	-0.139 (-2.24)	-0.437 (-5.14)	-0.211 (-4.81)	-0.206 (-4.74)	-0.225 (-4.57)	-0.103 (-1.60)	-0.101 (-1.57)	-0.367 (-4.58)
$0.5*\ln(x_4/x_1)^2$	-0.039 (-1.58)	0.002 (0.08)	-0.036 (-0.94)				-0.013 (-0.55)	0.023 (0.78)	-0.057 (-1.39)
$\ln(x_2/x_1)*\ln(x_3/x_1)$				-0.055 (-3.67)	-0.052 (-3.54)	-0.032 (-1.45)	-0.025 (-1.22)	-0.021 (-1.03)	-0.040 (-1.03)
$\ln(x_2/x_1)*\ln(x_4/x_1)$							-0.032 (-2.02)	-0.031 (-1.93)	0.012 (0.44)
$\ln(x_3/x_1)*\ln(x_4/x_1)$	-0.030 (-0.93)	-0.064 (-1.98)	0.003 (0.08)				-0.028 (-0.90)	-0.059 (-1.88)	0.061 (1.31)
$\ln(y_1)*\ln(x_2/x_1)$				0.086 (3.04)	0.098 (3.49)	0.006 (0.15)	0.046 (1.21)	0.064 (1.69)	-0.066 (-1.09)
$\ln(y_1)*\ln(x_3/x_1)$	0.229 (3.03)	0.241 (3.12)	0.079 (0.90)	-0.01 (-0.17)	0.010 (0.19)	-0.120 (-2.43)	0.214 (2.88)	0.221 (2.95)	0.007 (0.08)
$\ln(y_1)*\ln(x_4/x_1)$	-0.068 (-1.28)	-0.106 (-1.88)	-0.087 (-1.33)				-0.106 (-2.03)	-0.144 (-2.55)	-0.031 (-0.47)
$\ln(y_2)*\ln(x_2/x_1)$				0.025 (1.55)	0.015 (0.97)	0.107 (4.88)	0.013 (0.52)	0.000 (0.01)	0.101 (2.44)
$\ln(y_2)*\ln(x_3/x_1)$	-0.027 (-0.44)	-0.056 (-0.92)	0.169 (2.44)	0.201 (4.93)	0.171 (4.67)	0.171 (5.85)	-0.03 (-0.51)	-0.052 (-0.89)	0.224 (3.45)
$\ln(y_2)*\ln(x_4/x_1)$	0.043 (1.06)	0.102 (2.27)	0.047 (0.93)				0.067 (1.69)	0.121 (2.69)	-0.025 (-0.51)
$\ln(y_3)*\ln(x_2/x_1)$				-0.081 (-4.87)	-0.082 (-5.01)	-0.048 (-2.14)	-0.026 (-1.28)	-0.031 (-1.59)	0.017 (0.56)
$\ln(y_3)*\ln(x_3/x_1)$	-0.148 (-3.62)	-0.153 (-3.74)	-0.157 (-3.22)	-0.134 (-4.25)	-0.120 (-3.76)	-0.039 (-1.33)	-0.147 (-3.69)	-0.150 (-3.74)	-0.176 (-3.49)
$\ln(y_3)*\ln(x_4/x_1)$	0.044 (1.57)	0.037 (1.29)	0.090 (2.29)				0.054 (1.94)	0.052 (1.81)	0.090 (2.29)
t	0.015 (4.84)	0.016 (4.59)	0.010 (2.74)	0.003 (0.74)	0.007 (2.02)	0.000 (-0.14)	0.013 (3.96)	0.014 (3.81)	0.008 (2.10)
$0.5*t^2$	-0.006 (-5.91)	-0.006 (-5.59)	-0.006 (-2.79)	-0.009 (-8.55)	-0.009 (-8.52)	-0.012 (-7.34)	-0.009 (-8.22)	-0.009 (-7.76)	-0.011 (-5.24)
$t*\ln(y_1)$	0.006 (0.90)	0.000 (0.05)	0.003 (0.28)	0.012 (2.56)	0.010 (1.91)	0.012 (1.66)	0.01 (1.45)	0.003 (0.50)	0.019 (1.76)
$t*\ln(y_2)$	-0.011 (-2.03)	-0.008 (-1.51)	-0.007 (-0.83)	-0.02 (-5.57)	-0.017 (-4.35)	-0.023 (-4.45)	-0.014 (-2.47)	-0.010 (-1.74)	-0.016 (-1.85)
$t*\ln(y_3)$	0.005 (1.72)	0.009 (2.72)	0.004 (0.81)	0.005 (1.77)	0.005 (1.74)	0.008 (2.04)	0.002 (0.60)	0.005 (1.60)	-0.007 (-1.26)
$t*\ln(x_2/x_1)$				0.002 (0.63)	-0.001 (-0.52)	0.004 (1.10)	0.003 (0.99)	0.001 (0.33)	0.001 (0.26)
$t*\ln(x_3/x_1)$	0.010 (1.74)	0.014 (2.55)	-0.001 (-0.15)	0.021 (4.21)	0.025 (5.70)	0.017 (3.25)	0.009 (1.46)	0.013 (2.17)	-0.003 (-0.39)
$t*\ln(x_4/x_1)$	0.010 (2.92)	0.008 (2.43)	0.012 (1.89)				0.008 (2.28)	0.006 (1.86)	0.010 (1.60)
z_1 (rain)		0.000 (-0.15)			0.000 (-0.79)			0.000 (-0.41)	
z_2 (tmax)		-0.010 (-2.26)			-0.010 (-3.05)			-0.010 (-2.20)	
z_3 (tmin)		-0.004 (-1.56)			-0.004 (-1.62)			-0.004 (-1.53)	
δ_0			-0.069 (-0.52)			-0.464 (-1.90)			-0.323 (-1.31)
w_1 (rain)			0.000 (2.89)			0.000 (2.54)			0.000 (1.95)
w_2 (tmax)			0.040 (2.81)			0.079 (3.01)			0.070 (1.84)
w_3 (tmin)			0.019 (3.61)			0.030 (3.58)			0.028 (4.22)
γ	0.988 (80.35)	0.978 (63.35)	0.893 (21.88)	0.968 (44.66)	0.957 (40.39)	0.863 (20.18)	0.984 (82.07)	0.974 (63.04)	0.837 (6.95)
LLF	422.6	422.1	128.2	530.8	533.8	226.2	450.6	448.4	203.6
N° of observations	535	520	520	795	776	776	535	520	520

Case A: without weather, Case B: weather in production function and Case C: weather in inefficiency

Gamma (γ)³⁶, which explains the contribution of the inefficiency component on the variation of the composite error term, is on average 0.917, 0.918 and 0.846 for Case A, Case B and Case C models respectively. These values are higher for cost-quality models than for cost only models, which means that the fact that firms are not fully efficient is more completely explained when quality variables are introduced to the models. On average for these models, γ is equal to 0.974, 0.965 and 0.876 respectively; all of them are statistically significant at 1 per cent. In terms of the fit of the weather models, results from likelihood ratio tests indicate that Case A models are rejected in most cases in favour of Case B models³⁷ and fully rejected in favour of Case C models³⁸ at 1 per cent. This suggests that the effect of weather should not be ignored. This result is also in line with that from Jamasb *et al.* (2012), who suggest including weather due to its significant statistically effect in the cost function. . Coelli *et al.* (1999) also find similar results using the likelihood ratio test. The non-inclusion of weather is rejected in favour of models that include weather in the production function (input distance function).

Regarding Case B and Case C models, it is also convenient to analyse and compare both approaches in order to identify the approach that provides the best fit. The selection between Case B and Case C models appears to be a difficult task. In order to determine the best approach - i.e. weather in the production function versus weather in inefficiency - a set of nested models was built for doing proper comparisons. It is noteworthy that Case B and Case C models are not nested in each other. Two models are nested when one model is an extension of the other one. Nested models are used as an artifice for comparing models using specific tests such as log likelihood ratio (LLR) test. Following Coelli *et al.* (1999), the artificial nested models are built including weather variables in production and also in inefficiency as explanatory variables of technical efficiency. The idea is to test a null hypothesis using a likelihood test between (1) the nested models and the Case B models and (2) the nested models and the Case C models. The tests indicate that all models from Case B cannot be rejected in favour of the nested models. The tests also suggest that models from Case C are rejected in favour of the nested models at 1 per cent. This implies that models in which weather directly affects the production function are preferred due to the better fitness to the sample data. The following section discusses the results based on Case B models.

³⁶ $\gamma = \frac{\sigma_u^2}{\sigma^2}$, where $\sigma^2 = \sigma_u^2 + \sigma_v^2$

³⁷ Model 1 and Model 2 from Case A are rejected at 1 per cent and Model 3 at 10 per cent. Model 4, Model 5 and Model 6 cannot be rejected. However a Wald test suggests that Model 4 and Model 6 are rejected at 5 per cent and Model 5 at 1 per cent.

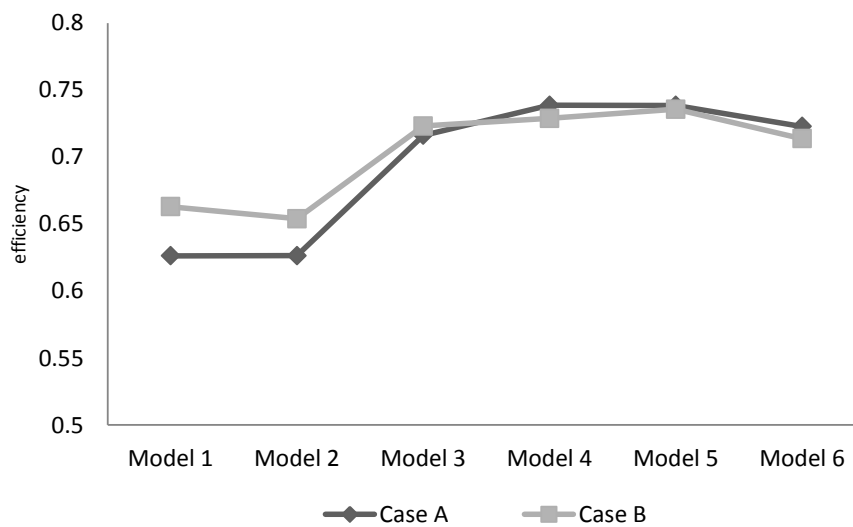
³⁸ In this case, all models from Case A are rejected in favour of models from Case C at 1 per cent.

6.2 Technical Efficiency

6.2.1 Global Effect

The global effect is measured by the variation of technical efficiency when comparing Case A models (without weather) with Case B models (with weather). In line with the discussion above we drop discussion of Case C. Figure 3 depicts the efficiency score for Case A models and Case B models (weather in input distance function).

Figure 3: Efficiency comparison (with and without weather)



Efficiency scores from Case B models vary from 0.663 to 0.714 and are on average slightly higher than those from Case A models. In addition, it can be seen that the influence of weather is more significant for cost only models than for cost-quality models³⁹. Efficiency scores increase on average by 5.12 per cent regarding cost only models. For cost-quality models, the variation is very low around -0.51 per cent. These results suggest that to some extent the inclusion of quality adjusts for weather, however, this does not necessarily justify regulators ignoring weather and only looking at quality. It is important to note that these results refer to average results. A company-level analysis may suggest that the variation is important for some countries and for some firms. Compa

This study's results are also in line with those from Growitsch *et al.* (2010). Their results suggest an increase of average efficiency by more than ten percentage points when weather

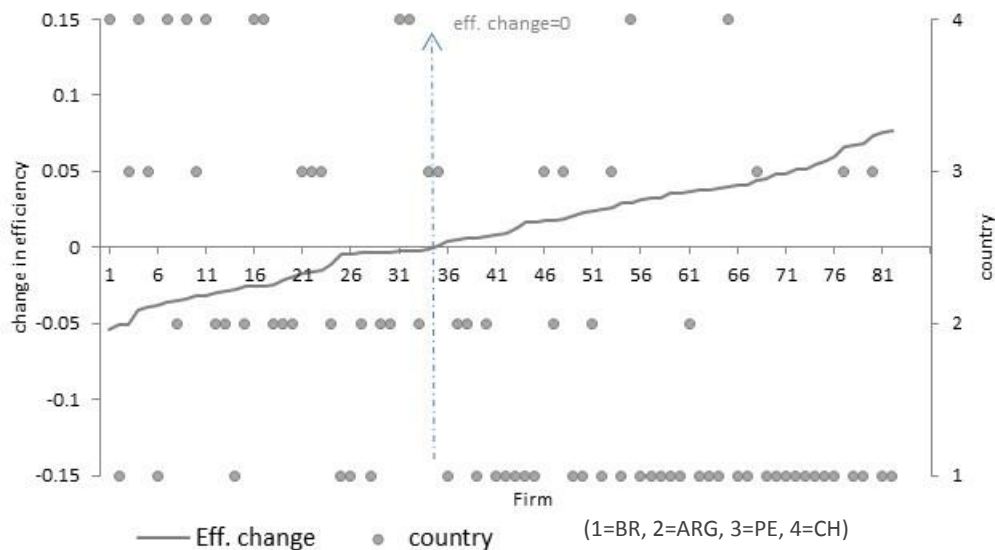
³⁹ The ANOVA test suggests that the variation on companies' efficiency (when weather variables are added) is statistically significant at 1 per cent only for Model 1 and Model 2.

variables and geographic factors are taken into consideration. In this study and for cost only models the increase of average efficiency exceeds 5 per cent. When comparing this study's results with those from Yu *et. al* (2009a), variables such as energy losses and customer minutes lost are those that have most internalised the effect of weather on efficiency, instead of the network length suggested by Yu *et al.* (2009a). From this study's results, the authors note that the inclusion or exclusion of network length does not affect the influence that weather has on efficiency. The increase in efficiency remains almost the same across models.

6.2.2 Company-level Effect

Figure 4 summarises the efficiency change across the six models when weather is added⁴⁰. On average a total of 48 firms increase their efficiency when weather is introduced with an average upward increase of 3.4 percentage points.

Figure 4: Average change of efficiency (1998-2008) when weather variables are introduced



The number of firms that decrease their efficiency is 34 with an average downward movement of 2.2 percentage points. Brazil is the country with the highest percentage number of firms that increase their efficiency, representing 84.6 per cent, followed by Peru with 50 per cent, Argentina with 33.3 per cent and Chile with 18.2 per cent. In addition, firms from Brazil and Peru are ones with the highest increase in efficiency (3.7 and 3.5 percentage points respectively). These results indicate that in general, firms from Brazil and Peru operate in less favourable weather conditions than companies based in Argentina or Chile.

⁴⁰ The location of the firms (which increase or decrease their efficiency) is indicated by the 'circle' marker. This takes the value of 1 for Brazilian firms, 2 for Argentine firms, 3 for Peruvian firms and 4 for Chilean firms.

In addition to the efficiency change, a ranking analysis at firm-level for each model indicates that firms from Brazil and Peru tend to increase their ranking in comparison with firms from Argentina and Chile. This result is also in agreement with that related to efficiency change, in which firms from Brazil and Peru have the highest percentage of firms that raise their efficiency when weather is introduced. Table 8 illustrates the number of firms that increase (↑), decrease (↓) or remain in the same position (≈) for each country and across models.

Table 8: Ranking variation per type of model

Country	Ranking variation - Number of firms per model																	
	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	↑	↓	≈	↑	↓	≈	↑	↓	≈	↑	↓	≈	↑	↓	≈	↑	↓	≈
Brazil	26	9	4	27	10	2	22	9	8	18	14	7	24	10	5	18	12	9
Argentina	1	17	0	1	16	1	6	10	2	0	8	1	9	8	1	3	6	0
Peru	6	7	1	7	5	2	4	8	2	5	8	1	5	9	0	4	8	2
Chile	2	9	0	1	9	1	1	9	1				1	10	0			
Total	35	42	5	36	40	6	33	36	13	23	30	9	39	37	6	25	26	11

↑: upper position, ↓: lower position, ≈: no variation

It can also be observed that depending on the models, the change of rankings for some specific firms is very notable with the maximum upward change of 28 positions and the maximum downward change of 24 positions, both related to Model 5. Thus in general on average across models (cost and cost-quality models) the introduction of weather produces important ranking changes for some firms. In summary, 37 firms improve, 38 firms worsen and 7 firms remain in the same position. The ranking variation differs across countries, where firms from Brazil are those with the highest number of firms that increase their rankings (61.5 per cent), followed by Peru (50 per cent), Argentina (22.2 per cent) and Chile (18.2 per cent)⁴¹. On average, the number of firms that change their ranking by more than 10 positions is 14 with a maximum upward change of 16 positions and the maximum downward change of 16 positions as well.

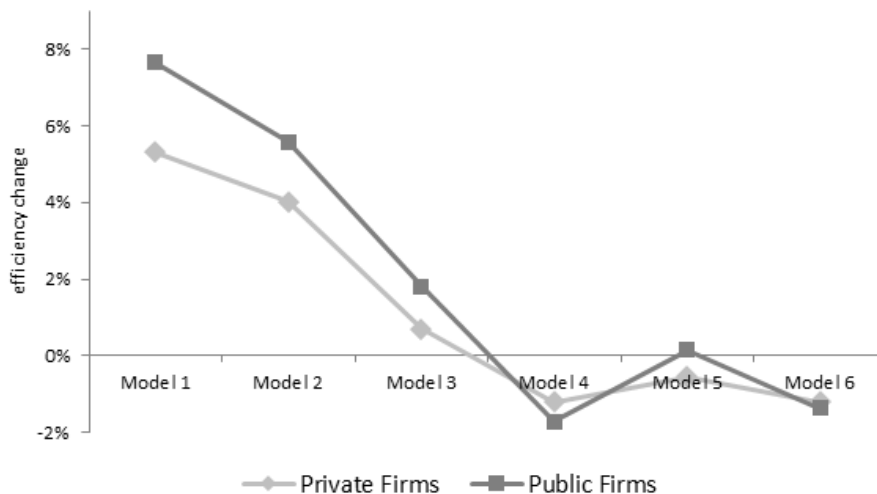
Following Coelli *et al.* (2003), when specific firms face important variations in efficiency due to the introduction of environmental variables, we recommend that regulators invite these specific firms to make a case for the appropriate adjustment of their efficiency scores. The same criteria should be applied when important ranking variations are observed.

⁴¹ Similar to the discussion regarding the change in efficiency, these figures represent the percentage of firms for each country.

6.2.3 Ownership-level Effect

Figure 5 shows the average efficiency change for private and public firms for the period 1998-2008. There are a total of 60 private and 22 public firms respectively. The introduction of weather produces a higher increase in efficiency in public firms than in private firms, for Model 1 and Model 2 (cost only models) and Model 3 (a cost-quality model). The difference of efficiency change between public and private firms is 2.6, 1.3 and 1.1 percentage points respectively⁴². These results suggest that on average, public firms operate in less favourable conditions than private ones.

Figure 5: Average change of efficiency following introduction of weather variables
per type of ownership



For the remaining models, in general private and public firms decrease their efficiency when weather is added. This again suggests that firms have adapted their networks based on the environmental conditions where they operate, thus the effect of weather on efficiency is much lower. However, the decrease in efficiency could be a sign of higher costs (inclusive of capex) in order to achieve the level of quality indicators established by regulators. For instance in Model 4, the decrease in efficiency is higher for public firms than for private firms. This could suggest that public firms incur higher operating costs when quality issues such as losses and customer hours lost are taken into consideration.

⁴² Based on the ANOVA results, the difference is significant for private firms in Model 1 and Model 2 at 1 per cent and for public firms in Model 1 and Model 2, but at 5 per cent.

In Model 6, the trend is the same but the gap between private and public efficiency change is much lower. Consequently, when total costs are taken into account and the full quality variables are included, the influence of weather appears to be nearly the same for both public and private firms. In Model 5 an opposite trend is observed. Public firms increase their efficiency very slightly and the private ones decrease their efficiency when weather is introduced. This also suggests that private firms could be operating in better weather conditions than the public ones.

In summary, it has been shown that on average public firms operate in worse weather conditions than private firms. This result makes sense for two related reasons. Firstly, less developed areas are affected by worse weather conditions in comparison with the most developed areas. Secondly, firms that operate in less developed areas are more likely to be publicly owned. A brief analysis regarding the Brazilian firms suggests that 50 per cent of the publicly owned firms (five out of 10) operate in areas with a human development index (HDI)⁴³ lower than the nationwide average. The average human development index of the five firms is 0.62 and the nationwide average is 0.69⁴⁴. In addition, for these five firms, four of them operate in areas with total annual rainfall much higher than the nationwide average. Regarding firms from Peru, a similar picture is observed. Figures indicate that 80 per cent of the publicly owned firms (eight out of 10) operate in areas with a low HDI. The average HDI regarding the eight firms is 0.56 and the nationwide average is 0.59⁴⁵. In terms of weather, of these eight firms, six operate in areas with annual total rainfall around 1,080 mm, against a nationwide average of 542 mm⁴⁶.

7. Conclusions

In this study, technical efficiency was estimated for 82 electricity distribution firms that operate in South America over the period 1998-2008. A stochastic frontier approach was selected and weather variables were included in the analysis. Different approaches to

⁴³ The Human Development Index is a composite statistic that assesses the development of a country (social and economic). The HDI is also available at state, city, district and municipal level, and other geographical divisions. It takes value from 0 to 1. The HDI was established in 1990 by the United Nations.

⁴⁴ PNUD, Atlas do Desenvolvimento Humano no Brasil, 2000 figures. Found at <http://www.pnud.org.br/atlas>.

⁴⁵ PNUD, Índice sobre desarrollo humano, Perú 2005. Found at <http://hdr.undp.org/sites/default/files/indh2009-peru-vol2-anexos.pdf>.

⁴⁶ Unfortunately, due to lack of information regarding HDI, it was not possible to determine statistically the significance of the previous results. This analysis requires the identification of HDI at company level. Thus, data of HDI at municipal, department and district level is mandatory. For example in the case of Brazil, the HDI at municipal level (IDH-M) was only available for two years: 1990 and 2000. In the case of Peru HDI is available at district level (in most cases from 2002 onwards) however the information is in pdf format.

determining the preferred weather models were analysed, namely the inclusion of weather in the production function vs its inclusion in the inefficiency term. Based on statistical hypothesis testing and nested models, we conclude that the inclusion of weather in the production function should be the preferred approach. Geographic information systems made it possible to locate the firms' service areas and to allocate their respective meteorological stations and NASA coordinates (flash rate).

The stochastic frontier analysis shows that across the models examined a large number of firms exist where weather affects the measured efficiency. This to some extent reflects the appropriate selection of weather variables due to the large effect they can have. In comparison with similar studies which incorporate weather into efficiency analysis (Korhonen and Syrjänen, 2003; Yu *et al.*, 2009a; Nillesen and Pollitt, 2010; Jamasb *et al.*, 2012; Growitsch *et al.*, 2010) this paper shows that the impact of weather is quite subtle and depends on the nature of the sample.

In summary, our stochastic frontier analysis confirms that weather can have significant impacts on average measured efficiency; this effect is stronger in some countries than in others; and for specific firms weather affects their absolute and relative performance substantially. Our findings therefore highlight the importance of careful attention to the potential role for weather as a determinant of electricity distribution efficiency.

Acknowledgments

The authors are very grateful to the many energy regulators, electricity distribution companies, associations of electricity distribution firms, national statistical agencies, national meteorological agencies, national security and insurance agencies from Argentina, Brazil, Chile and Peru who helped them with data collection. In addition, the authors wish to acknowledge Steven Goodman (NASA) and Rachel Albrecht (Earth System Science Interdisciplinary Centre-University of Maryland) for their great support in the provision of satellite data, to Luis Orea for his fantastic help in managing satellite data using Gauss and to one anonymous referee for his valuable suggestions. This research was supported by RCUK – ESRC (Grant Number: RES-152-25-1002). The views expressed herein are those of the authors and do not reflect the views of EPRG.

References

- Aigner, D.J., Lovell, C.A.K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, Vol. 6, No. 1, pp. 21-37.
- ANEEL (2007). *Manual de Contabilidade do Serviço Público de Energia Elétrica*, Brasília: Agência Nacional de Energia Elétrica.
- Battese, G. and Coelli, T. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, Vol. 3, No. 1-2, pp. 153-169.
- Battese G. and Coelli, T. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, Vol. 20, No. 2, pp. 325-332.
- Bjørndal, M., Burheim, E.N; Dahl, E.I., Bjørndal, E. (2016). *Learning and adaptation under incentive regulation: A survey of Norwegian electricity distribution companies*, Norwegian School of Economics, BELEC, October 2016.
- BLS (2016). *CPI Detailed Report Data for December 2016*. U.S. Bureau of Labor Statistics, Washington DC.
- Campbell, Richard J. (2012). Weather-related power outages and electric system resiliency. *Congressional Research Service*, 7-5700 R42696, CRS Report for the Congress.
- CEIDS (2001). *The cost of power disturbances to industrial and digital economy economies*. Submitted to EPRI's Consortium for Electric Infrastructure for a Digital Society (CEIDS), June 2001, Madison, WI.
- Coelli, T., Perelman, S. and Romano, E. (1999). Accounting for environmental influences in stochastic frontier models: with application to international airlines. *Journal of Productivity Analysis*, Vol. 11, No. 3, pp. 251-273.
- Coelli, T., Estache, A., Perelman, S. and Trujillo, L. (2003). *A primer on efficiency measurement for utilities and transport regulators*. Washington, DC: The World Bank.
- Coelli, T., Prasada Rao, D.S. and Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Second Edition, New York: Springer Science and Business Media, Inc.
- Christensen, L.R., Jorgenson, D.W. and Lau, L.J. (1973). Transcendental Logarithmic Production Functions. *Review of Economics and Statistics*, Vol. 55, No. 1, pp. 28-45.
- Cuesta, Rafael A. (2000). A production model with firm-specific temporal variation in technical inefficiency: With application to Spanish dairy farms. *Journal of Productivity Analysis*, Vol. 13, No. 2, pp. 139-158.
- DOE (1981). *The national electric reliability study: technical study reports*, DOE/EP-0005. U.S Department of Energy, Washington DC.
- DOE (2016). *Smart Grid Investment Grant Program Final Report*. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability, Dec. 2016, Washington DC.
- Desli, E., Ray, S.C., and Kumbhakar, S.C. (2003). A dynamic stochastic frontier production model with time-varying efficiency. *Applied Economics Letters*, 10, pp. 623–626.

EI (2012). Out of Sight, Out of Mind 2012. An Updated Study on the Undergrounding of Overhead Power Lines. *Edison Electric Institute*, Washington D.C.

EIA (2009). *Table CE3.2 Household Site End-Use Consumption in the Northeast Region, Totals and Averages, 2009*. U.S. Energy Information Administration.

Estache A., Rossi M.A. and Ruzzier, C.A. (2004). The case for international coordination of electricity regulation: Evidence from the measurement of efficiency in South America. *Journal of Regulatory Economics*, Vol. 25, No. 3, pp. 271-295.

EU (2016). An EU Strategy for Heating and Cooling, SWD (2016) 24 final, Part 1/2. European Commission, Brussels.

Fan, Y., Li, Q., and Weersink, A. (1996). Semiparametric estimation of stochastic production frontier models. *Journal of Business and Economic Statistics*, Vol. 14 (4), pp. 460–468.

Giannakis, D., Jamasb, T. and Pollitt, M.G. (2005). Benchmarking and incentive regulation of quality of service: An application to the UK electricity distribution networks. *Energy Policy*, Vol. 33, No. 1, pp. 2256-2271.

Gönen, T. (2007). *Electric Power Distribution System Engineering*, Second Edition, New York: CRC Press.

Greene, W.H. (2004). Distinguishing between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Care Systems. *Health Economics*, Vol. 13, No. 10, pp. 959–980.

Greene, W. H. (2005). Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model. *Journal of Econometrics*, Vol. 126, No. 2, pp. 269–303.

Growitsch, C., Jamasb, T. and Wetzel, H. (2010). Efficiency effects of quality of service and environmental factors: Experience from Norwegian Electricity Distribution. *EPRG Working Paper 1025*, Cambridge Working Paper in Economics 1050. University of Cambridge.

Hines, P., Apt, J., Talukdar, S. (2008). Trends in the History of Large Blackouts in the United States. *Proc. Of the IEEE Power and Energy Society*, 2008.

Horrace, W.C. and Parmeter, C.F. (2011), Semiparametric deconvolution with unknown error variance, *Journal of Productivity Analysis*, 35, pp. 129-141.

Jamasb, T., Orea, L. and Pollitt, M.G. (2012). Estimating the marginal cost of quality improvements: The case of the UK electricity distribution companies. *Energy Economics*, Vol. 34, pp. 1498-1506.

Keener, Ronald N. Jr. (1997). *The impact of weather on daily electric utility operations*. Workshop on the social and economics impact of weather, CO, April 2-4, 1997.

Korhonen, P.J. and Syrjanen, P.J. (2003). Evaluation of Cost Efficiency in Finnish Electricity Distribution. *Annals of Operations Research*, Vol. 121, No. 1-4, pp. 105-122.

Kumbhakar, S., Ghosh, S. and McGucking, J.T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. Dairy Firms. *Journal of Business and Economic Studies*, Vol. 9, No. 3, pp. 279-286.

Kumbhakar, S. and Knox Lovell C.A. (2000). *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Kumbhakar, S.C., Park, B.U., Simar, L., and Tsionas, E.G. (2007). Non parametric stochastic frontiers: A local maximum likelihood approach. *Journal of Econometrics*, 137(1), pp. 1–27.

- Kumbhakar, S. (2013). Specification and estimation of multiple output technologies: A primal approach. *European Journal of Operational Research*, Vol. 217, No. 3, pp. 465-473.
- Kuosmanen, T. and Johnson A.L. (2010). Data envelopment analysis as nonparametric least-squares regression, *Operations Research*, Vol. 58 (1), 149-160.
- Kuosmanen, T. and Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. *Journal of Productivity Analysis*, Vol. 38 (1), pp. 11-28.
- Lee, Y.H. (2010). Group-specific stochastic production frontier models with parametric specifications. *European Journal of Operational Research*, Vol. 200, No. 2, pp. 508-517.
- Llorca, M., Orea, L. and Pollitt, M. (2014). Using latent class approach to cluster firms in benchmarking: An application to the US electricity transmission industry. *Operations Research Perspectives*, Vol. 1, pp. 6-17.
- Llorca, M., Orea, L. and Pollitt, M. (2016). Efficiency and environmental factors in the US electricity transmission industry. *Energy Economics*, Vol. 55, pp. 234-246.
- Meeusen W. and van den Broeck, J. (1977). Efficiency estimation from Coob-Douglas production functions with composed error. *International Economic Review*, Vol. 18, No. 2, pp. 435-444.
- Migueis, V.L., Camanho, A.S., Bjørndal, E., Bjørndal, M. (2012). Productivity change and innovation in Norwegian electricity distribution companies. *Journal of Operational Research Society*, Vol. 63, pp. 982-990.
- Mills, E. (2012). *Electric grid disruptions and extreme weather*. Lawrence Berkeley National Laboratory. Presented at US Disaster Reanalysis Workshop, National Climatic Data Center, Asheville, NC.
- MINEM (1994). *Manual de costos para empresas de electricidad concesionarias y/o autorizadas*. Resolución Ministerial No. 197-94-EM/VME, Lima, Perú: Ministerio de Energía y Minas.
- Mota, R.L. (2004). Comparing Brazil and USA electricity distribution performance: What was the impact of privatization?. *Cambridge Working Paper in Economics CWPE 0423, CMI Working Paper 39*. University of Cambridge.
- Nillesen, P. and Pollitt, M.G. (2010). Using regulatory benchmarking techniques to set company performance targets: the case of US electricity. *Competition and Regulation in Network Industries*, Vol. 11, No. 1, pp. 50-84.
- NordREG (2011). *Economic regulation of electric grids in Nordic countries, Report 7/2011*. Nordic Energy Regulators, Denmark, December 2011.
- OSINERGMIN (2003). *Estadísticas sobre la calidad del servicio eléctrico*. Lima, Perú: Organismo Supervisor de la Inversión en Energía y Minería.
- Pabla, A.S. (2005). *Electric power distribution*, New York: McGraw-Hill.
- PCEA-DOE (2013). *Economics benefits of increasing electric grid resilience to weather outages*. Executive Office of the President. Prepared by the President's Council of Economic Advisers and the U.S. Department of Energy's Office of Electricity Delivery and Energy Reliability, Washington DC.
- Pinto, Osmar Jr. and Almeida Pinto, Iara (2008). *Relâmpagos*. 2ª edição, Sao Paulo: Editora Brasiliense.
- Pollitt, M. G. (1995). *Ownership and performance in electric utilities*. Oxford: Oxford University Press.
- Rossi, M. (2007). Labour productivity change estimates as an input for X-factors, in Chisari, O. (ed.) *Regulatory Economics and Quantitative Methods: Evidence from Latin America*. The CRC Series on Competition, Regulation and Development. Cheltenham, UK: Edward Elgar Publishing Limited.

Rudnick, H., Barroso, L., Cunha, G., Mocarquer, S. (2014). Electricity-Gas Integration Challenges in South America. *IEEE power & energy magazine*, October 2014.

SEC (2006). *Sistema de cuentas para informar y valorizar los costos e ingresos de explotación de las concesionarias del servicio público de distribución eléctrica*. Santiago de Chile: Superintendencia de Electricidad y Combustibles.

Secretaría de Energía (2008). *Informe estadístico del sector eléctrico 2008*. Buenos Aires: Secretaria de Energía del Ministerio de Planificación Federal, Administración Pública y Servicios.

Shephard, R.W. (1970). *Theory of cost and production functions*. Princeton: Princeton University Press.

Simar, L. and Wilson, P. (2007). Estimation and Inference in Two-stage, Semi-parametric Models of Production Processes. *Journal of Econometrics*, Vol. 136, No. 1, pp. 31-64.

Simar, L., Van Keilegom, I., and Zelenyuk, V. (2016). Nonparametric least square methods for stochastic frontier models. *Journal of Productivity Analysis*, doi:10.1007/s11123-016-0474-2.

Short, Tom A. (2006). *Distribution reliability and power quality*, Second Edition, Boca Raton: CRC Press and Taylor & Francis Group.

Sullivan, M.J., Mercurio, M., Schellenberg, J. (2009). *Estimated value of service reliability for electric utility customers in the United States*. Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA.

Sun, K., and Kumbhakar, S.C. (2013). Semiparametric smooth-coefficient stochastic frontier model. *Economic Letters*, 120 (2), pp. 305–309.

Sun, K., Kumbhakar, S.C., and Tveteras, R. (2015). Productivity and efficiency estimation: A semiparametric stochastic cost frontier analysis. *European Journal of Operational Research*, 245, pp. 194-203.

Tovar, B., Ramos-Real, F.J. and Fagundes de Almeida, E. (2011). Firm size and productivity. Evidence from the electricity distribution industry in Brazil. *Energy Policy*, Vol. 39, No. 2, pp. 826-833.

Tsionas, E.G. (2006). Interference in dynamic stochastic frontier models. *Journal of Applied Econometrics*, 21, pp. 669-676.

Tubelo, Renata .C.S., Rodrigues, Lucelia T., Gillott, M. (2014). A Comparative study of the Brazilian Energy labelling system and the Passivhaus Standard for Housing. *Buildings* 2014, 4, pp. 207-221.

World Bank (2008). *Benchmarking data on the electricity distribution sector in the Latin American and the Caribbean Region 1995-2005*. Washington, DC: The World Bank.

Yang, H.L. and Pollitt, M.G. (2009). Incorporating both undesirable outputs and uncontrollable variables into DEA: the performance of Chinese coal-fired power plants. *European Journal of Operational Research*, Vol. 197, No. 3, pp. 1095-1105.

Yu, W., Jamasb, T. and Pollitt, M.G. (2009a). Does weather explain the cost and quality performance? An analysis of UK electricity distribution companies. *Energy Policy*, Vol. 37, No. 11, pp. 4177-4188.

Yu, W., Jamasb, T. and Pollitt, M.G. (2009b). Willingness-to-Pay for Quality of Service: An Application to Efficiency Analysis of the UK Electricity Distribution Utilities. *The Energy Journal*, Vol. 30, No. 4, pp. 1-48.

Zhang, R., Sun, K., Delgado, M.S., and Kumbhakar, S.C. (2012). Productivity in China's high technology industry: Regional heterogeneity and R&D. *Technological Forecasting and Social Change*, 79, pp. 127–141.

Zhang, Q., Xu, Z., Feng, T., and Jiao, J. (2015). A dynamic stochastic frontier model to evaluate regional financial efficiency: Evidence from Chinese country-level panel data. *European Journal of Operational Research*, 241, pp. 907-916.

Appendix 1: Weather Variables

Following Table A1, all the weather parameters have the correct sign; the worse the weather conditions, the higher the costs. It can be seen that the effect of weather variables depends on the scenario and models. In the first scenario, the most significant variable is the maximum absolute temperature but only for cost-quality models.

Table A1: Weather estimates per model⁴⁷

Model	Scenario 1					Scenario 2				Scenario 3			
	rain	tmax	tmin	fr	γ	rain	tmax	tmin	γ	fr	tmax	tmin	γ
M1	-0.000*	-0.005	-0.003	-0.003	0.84	0	-0.004	-0.007***	0.84	-0.005	-0.005	-0	0.83
M2	0	-0.005	-0.002	-0.002	0.80	0	-0.005	-0.005**	0.80	-0.004	-0.005	-0	0.80
M3	-0.000**	-0.011***	-0.003	-0.003	0.96	-0.000*	-0.009***	-0.005**	0.95	-0.005*	-0.010***	-0	0.96
M4	0	-0.011**	-0.005*	-0.004	0.98	0	-0.010**	-0.004	0.98	-0.003	-0.009**	-0	0.98
M5	0	-0.012***	-0.002	-0.003	0.97	0	-0.010***	-0.004	0.96	-0.004	-0.011***	-0	0.98
M6	0	-0.010**	-0.004	-0.003	0.98	0	-0.010**	-0.004	0.97	-0.003	-0.008*	-0	0.98

Significance levels: * p<0.1; ** p<0.05; *** p<0.01

In cost only models, total rainfall is statistically significant but weak due to the very low coefficient. The minimum absolute temperature is only statistically significant at 10 per cent in Model 4. Flash rate is not statistically significant across models. Under the second scenario, we have a similar picture to the previous one; however, in this case the minimum absolute temperature is statistically significant in relation to the cost only models. An increase in the coefficient is also observed on average in absolute values. The coefficients of the maximum absolute temperature remain the same in general. Regarding the last scenario, a similar trend as seen in the first scenario is observed. Maximum absolute temperature continues being the weather variable that influences efficiency the most when quality variables are taken into account. Flash rate is only significant in Model 3. In terms of cost only models, none of the weather variables is statistically significant. In terms of the inefficiency term, we observe that the gamma values indicate that firms are not fully efficient when weather is introduced across the three scenarios. Gamma varies from 0.80 to 0.98 and is statistically significant at the 1 per cent level.

⁴⁷ A total of 18 models were analysed. This is a simplified table that only shows the weather coefficients, their respective significance level and gamma values.